

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 12:23:41 ; Search time 2617 Seconds
(without alignments)
2568.876 Million cell updates/sec

Title: US-09-673-716-1
Perfect score: 231
Sequence: 1 ggggggtcgcaggctaagca.....ttcaccatgaggtgtgctt 231

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rod.*
- 36: em.htg_nam.*
- 37: em.htg_vrt.*
- 38: em.sy.*
- 39: em.higo_hum.*
- 40: em.higo_mus.*
- 41: em.higo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	DB ID	Description
1	231	100.0	1086	10	AF250998 Mus muscu
2	231	100.0	1086	10	AF250999 Mus muscu
3	231	100.0	2701	10	BC020078 Mus muscu
4	231	100.0	42511	2	AC100548 Mus muscu
5	231	100.0	47266	2	AC129555 Mus muscu
6	231	100.0	59079	2	AC116705 Mus muscu
7	231	100.0	60388	2	AC124106 Mus muscu
8	231	100.0	61926	2	AC113184 Mus muscu
9	231	100.0	64042	2	AC101394 Mus muscu
10	231	100.0	66393	2	AC100424 Mus muscu
11	231	100.0	66489	2	AC101278 Mus muscu
12	231	100.0	66489	2	AC101278 Mus muscu
13	231	100.0	69090	2	AC105951 Mus muscu
14	231	100.0	70360	2	AC101122 Mus muscu
15	231	100.0	71466	2	AC123753 Mus muscu
16	231	100.0	84250	10	AF481949 Mus muscu
17	231	100.0	98653	10	AL731664 Mouse DNA
18	231	100.0	121617	2	AL670597 Mus muscu
19	231	100.0	126321	2	AC102180 Mus muscu
20	231	100.0	134025	2	AC107697 Mus muscu
21	231	100.0	134686	10	AL626805 Mouse DNA
22	231	100.0	141025	10	AL807250 Mouse DNA
23	231	100.0	150017	2	AL773509 Mus muscu
24	231	100.0	150017	2	AL844603 Mus muscu
25	231	100.0	152040	2	AC127577 Mus muscu
26	231	100.0	162681	10	AC122059 Mus muscu
27	231	100.0	165015	2	AC121307 Mus muscu
28	231	100.0	166764	2	AC115875 Mus muscu
29	231	100.0	171576	2	AC102493 Mus muscu
30	231	100.0	171826	2	AC119810 Mus muscu
31	231	100.0	172533	2	AC120551 Mus muscu
32	231	100.0	173417	10	AL607143 Mouse DNA
33	231	100.0	175213	10	AC079680 Mus muscu
34	231	100.0	177353	2	AC124471 Mus muscu
35	231	100.0	179043	2	AL808015 Mus muscu
36	231	100.0	179098	2	AC117690 Mus muscu
37	231	100.0	180634	2	AC102775 Mus muscu
38	231	100.0	180822	2	AC101850 Mus muscu
39	231	100.0	181596	2	AC117820 Mus muscu
40	231	100.0	182005	2	AC124140 Mus muscu
41	231	100.0	182882	2	AC101803 Mus muscu
42	231	100.0	183536	10	AC084416 Mus muscu
43	231	100.0	183816	2	AC102079 Mus muscu
44	231	100.0	184663	2	AL772299 Mus muscu
45	231	100.0	185691	2	AL844578 Mus muscu

ALIGNMENTS

RESULT 1

AF250998

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AF250998 1086 bp DNA linear ROD 16-MAY-2001
Mus musculus RTE-clone1 RNA transport element sequence.

AF250998.1 GI:14090507

Mus musculus.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 1086)

Nappi, F., Schneider, R., Zolotukhin, A., Smulevitch, S.,

Michalowski, D., Bear, J., Felber, B.K. and Pavlakis, G.N.

Identification of a novel posttranscriptional regulatory element by

using a rev- and RRE-mutated human immunodeficiency virus type 1
DNA proviral clone as a molecular trap
J. Virol. 75 (10), 4558-4569 (2001)
21211606
PUBMED
11312326
REFERENCE
2 (bases 1 to 1086)
Nappi, F., Schneider, R. and Pavlakis, G.N.
Direct Submission
Submitted (31-MAR-2000) Human Retrovirus Section, BRL, National
Cancer Institute-Frederick, Bldg 535, Room 226, Frederick, MD
21702, USA
FEATURES
source
1..1086
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RTE-clone1"
misc_feature
381..627
/note="fragment M1; contains RNA transport element (RTE);
posttranscriptional regulatory element"
/evidence=experimental
BASE COUNT
248 a 268 c 276 g 294 t
ORIGIN
Query Match 100.0%; Score 231; DB 10; Length 1086;
Best Local Similarity 100.0%; Pred. No. 3.7e-65;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGGGGTGGAGGCTAAGCACTGCACAGAGATAGCTTGTGTTGGCATCTCTGTGGAAGG 60
|||||
Db 393 GTGGGGTGGAGGCTAAGCACTGCACAGAGATAGCTTGTGTTGGCATCTCTGTGGAAGG 452
Qy 61 CACGCTCTGATTGTCATGAAGGTTTCAGTGTCTAGTTCCTTCCCGCCAGGAAACGACACG 120
|||||
Db 453 CACGCTCTGATTGTCATGAAGGTTTCAGTGTCTAGTTCCTTCCCGCCAGGAAACGACACG 512
Qy 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGGCCCT 180
|||||
Db 513 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGGCCCT 572
Qy 181 ATGCTTGACACTGGGATCAGACCTTACCTTCACCCATGAGGCTTGCTT 231
|||||
Db 573 ATGCTTGACACTGGGATCAGACCTTACCTTCACCCATGAGGCTTGCTT 623
RESULT 2
AF250999
LOCUS
Mus musculus RTE-clone3 RNA transport element sequence.
AF250999 1086 bp DNA linear ROD 16-MAY-2001
ACCESSION
AF250999
VERSION
AF250999.1 GI:14090508
KEYWORDS
Mus musculus.
SOURCE
Mus musculus.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1086)
Nappi, F., Schneider, R., Zolotukhin, A., Smulevitch, S.,
Michalowski, D., Bear, J., Feller, B.K. and Pavlakis, G.N.
IDENTIFICATION of a novel posttranscriptional regulatory element by
using a rev- and RRE-mutated human immunodeficiency virus type 1
DNA proviral clone as a molecular trap
J. Virol. 75 (10), 4558-4569 (2001)
21211606
PUBMED
11312326
REFERENCE
2 (bases 1 to 1086)
Nappi, F., Schneider, R. and Pavlakis, G.N.
Direct Submission
Submitted (31-MAR-2000) Human Retrovirus Section, BRL, National
Cancer Institute-Frederick, Bldg 535, Room 226, Frederick, MD
21702, USA
FEATURES
source
1..1086
/organism="Mus musculus"

/db_xref="taxon:10090"
/clone="RTE-clone3"
381..627
misc_feature
/note="fragment M1; contains RNA transport element (RTE);
posttranscriptional regulatory element"
BASE COUNT
247 a 268 c 277 g 294 t
ORIGIN
Query Match 100.0%; Score 231; DB 10; Length 1086;
Best Local Similarity 100.0%; Pred. No. 3.7e-65;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GTGGGGTGGAGGCTAAGCACTGCACAGAGATAGCTTGTGTTGGCATCTCTGTGGAAGG 60
|||||
Db 393 GTGGGGTGGAGGCTAAGCACTGCACAGAGATAGCTTGTGTTGGCATCTCTGTGGAAGG 452
Qy 61 CACGCTCTGATTGTCATGAAGGTTTCAGTGTCTAGTTCCTTCCCGCCAGGAAACGACACG 120
|||||
Db 453 CACGCTCTGATTGTCATGAAGGTTTCAGTGTCTAGTTCCTTCCCGCCAGGAAACGACACG 512
Qy 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGGCCCT 180
|||||
Db 513 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGGCCCT 572
Qy 181 ATGCTTGACACTGGGATCAGACCTTACCTTCACCCATGAGGCTTGCTT 231
|||||
Db 573 ATGCTTGACACTGGGATCAGACCTTACCTTCACCCATGAGGCTTGCTT 623
RESULT 3
BC020078
LOCUS
Mus musculus, Similar to RIKEN cdNA 1700066C05 gene, clone
MGC:28125 IMAGE:3980327, mRNA, complete cds.
ACCESSION
BC020078
VERSION
BC020078.1 GI:18043256
KEYWORDS
MGC.
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 2701)
Strausberg, R.
Direct Submission
Submitted (19-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgaphs-r@mail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunsaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 36 Row: e Column: 4
This clone was selected for full length sequencing because it
passed the following selection criteria: Similarity but not
identity to protein.
FEATURES
source
1..2701
/organism="Mus musculus"
/db_xref="taxon:10090"

* 17219 17318: gap of 100 bp in length
 * 17319 18010: contig of 692 bp in length
 * 18011 18110: gap of 100 bp in length
 * 18111 18790: contig of 680 bp in length
 * 18791 18890: gap of 100 bp in length
 * 18891 19594: contig of 704 bp in length
 * 19595 19694: gap of 100 bp in length
 * 19695 20385: contig of 688 bp in length
 * 20386 20482: gap of 100 bp in length
 * 20483 21165: contig of 683 bp in length
 * 21166 21265: gap of 100 bp in length
 * 21266 21972: contig of 707 bp in length
 * 21973 22072: gap of 100 bp in length
 * 22073 22753: contig of 681 bp in length
 * 22754 22853: gap of 100 bp in length
 * 22854 23527: contig of 674 bp in length
 * 23528 23627: gap of 100 bp in length
 * 23628 24334: contig of 707 bp in length
 * 24335 24434: gap of 100 bp in length
 * 24435 25132: contig of 698 bp in length
 * 25133 25232: gap of 100 bp in length
 * 25233 25910: contig of 678 bp in length
 * 25911 26010: gap of 100 bp in length
 * 26011 26691: contig of 681 bp in length
 * 26692 26791: gap of 100 bp in length
 * 26792 27469: contig of 678 bp in length
 * 27470 27569: gap of 100 bp in length
 * 27570 28256: contig of 687 bp in length
 * 28257 28356: gap of 100 bp in length
 * 28357 29071: contig of 715 bp in length
 * 29072 29171: gap of 100 bp in length
 * 29172 29868: contig of 697 bp in length
 * 29869 29968: gap of 100 bp in length
 * 29969 30719: contig of 751 bp in length
 * 30720 30819: gap of 100 bp in length
 * 30820 31527: contig of 708 bp in length
 * 31528 31627: gap of 100 bp in length
 * 31628 32310: contig of 683 bp in length
 * 32311 32410: gap of 100 bp in length
 * 32411 33099: contig of 689 bp in length
 * 33100 33199: gap of 100 bp in length
 * 33200 33877: contig of 678 bp in length
 * 33878 33977: gap of 100 bp in length
 * 33978 34666: contig of 689 bp in length
 * 34667 34766: gap of 100 bp in length
 * 34767 35460: contig of 694 bp in length
 * 35461 35560: gap of 100 bp in length
 * 35561 36271: contig of 711 bp in length
 * 36272 36371: gap of 100 bp in length
 * 36372 37043: contig of 674 bp in length
 * 37046 37145: gap of 100 bp in length
 * 37146 37835: contig of 690 bp in length
 * 37836 37935: gap of 100 bp in length
 * 37936 38609: contig of 674 bp in length
 * 38610 38709: gap of 100 bp in length
 * 38710 39386: contig of 677 bp in length
 * 39387 39486: gap of 100 bp in length
 * 39487 40166: contig of 680 bp in length
 * 40167 40266: gap of 100 bp in length
 * 40267 40937: contig of 671 bp in length
 * 40938 41037: gap of 100 bp in length
 * 41038 41717: contig of 680 bp in length
 * 41718 41817: gap of 100 bp in length
 * 41818 42511: contig of 694 bp in length.

FEATURES
 source

1. .42511
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="RP23-152017"
 /clone_lib="RPCI-23 Female Mouse BAC"

10831 a 7531 c 7714 g 10953 t 5482 others

BASE COUNT
 ORIGIN

Query Match 100.0%; Score 231; DB 2; Length 42511;
 Best Local Similarity 100.0%; Pred. No. 3.le-65;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGGGTCGAGGCTAAGCACTGCACAGAGATAGCTTGTGTGGCATCTGTGGAAGG 60
 Db 27042 GTGGGTCGAGGCTAAGCACTGCACAGAGATAGCTTGTGTGGCATCTGTGGAAGG 27101
 QY 61 CACGTCCTGATTGCATGAAGGTTCAAGTTCCTAGTTCCTCCCTCCCGAGGAAAAACACACG 120
 Db 27102 CACGTCCTGATTGCATGAAGGTTCAAGTTCCTAGTTCCTCCCTCCCGAGGAAAAACACACG 27161
 QY 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTAAAGGATGGTTTTGTAGGCCCTT 180
 Db 27162 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTAAAGGATGGTTTTGTAGGCCCTT 27221
 QY 181 ATGCTTGACACTGGGGATCAGACCTCTACCTTACCCATGAGGCTTCTT 231
 Db 27222 ATGCTTGACACTGGGGATCAGACCTCTACCTTACCCATGAGGCTTCTT 27272

RESULT 5
 AC129555/c
 LOCUS AC129555
 DEFINITION Mus musculus clone RP24-90F19, LOW-PASS SEQUENCE SAMPLING.
 AC129555
 VERSION AC129555.1 GI:22004399
 HTG: HTGS_PHASEO.
 SOURCE house mouse.
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 47266)
 Birren,B., Nusbaum,C. and Lander,E.
 Mus musculus chromosome, clone RP24-90F19
 Unpublished
 2 (bases 1 to 47266)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 AUTHORS
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
 Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L., Boukhgalter,B.,
 Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Faros,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
 Gardyna,S., Gord,S., Graham,L., Grand-pierre,N., Hagos,B.,
 Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
 Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
 Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C.,
 McCarthy,M., Meldrim,J., Menes,L., Mihova,T., Mienga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K.,
 Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
 Roman,J., Roy,A., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
 Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
 Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT
 Direct Submission
 Submitted (30-JUL-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L26515
 Center clone name: 90_F_19

 * NOTE: This record contains 59 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
674 773: contig of 673 bp in length
774 773: gap of 100 bp
1477 1576: contig of 703 bp in length
1577 2266: contig of 690 bp in length
2267 2366: gap of 100 bp
2367 3042: contig of 676 bp in length
3043 3142: gap of 100 bp
3143 3833: contig of 691 bp in length
3834 3933: gap of 100 bp
3934 4630: contig of 697 bp in length
4631 4730: gap of 100 bp
4731 5450: contig of 720 bp in length
5451 5550: gap of 100 bp
5551 6287: contig of 737 bp in length
6288 6387: gap of 100 bp
6388 7071: contig of 684 bp in length
7072 7171: gap of 100 bp
7172 7863: contig of 692 bp in length
7864 7963: gap of 100 bp
7964 8675: contig of 712 bp in length
8676 8775: gap of 100 bp
8776 9495: contig of 720 bp in length
9496 9595: gap of 100 bp
9596 10331: contig of 736 bp in length
10332 10431: gap of 100 bp
10432 11171: contig of 740 bp in length
11172 11271: gap of 100 bp
11272 11951: contig of 680 bp in length
11952 12051: gap of 100 bp
12052 12737: contig of 686 bp in length
12738 12837: gap of 100 bp
12838 13514: contig of 677 bp in length
13515 13614: gap of 100 bp
13615 14314: contig of 700 bp in length
14315 14414: gap of 100 bp
14415 15115: contig of 701 bp in length
15116 15215: gap of 100 bp
15216 15927: contig of 712 bp in length
15928 16027: gap of 100 bp
16028 16718: contig of 691 bp in length
16719 16818: gap of 100 bp
16819 17536: contig of 718 bp in length
17537 17636: gap of 100 bp
17637 18315: contig of 679 bp in length
18316 18415: gap of 100 bp
18416 19108: contig of 693 bp in length
19109 19208: gap of 100 bp
19209 19902: contig of 694 bp in length
19903 20002: gap of 100 bp
20003 20703: contig of 701 bp in length
20704 20803: gap of 100 bp
20804 21512: contig of 709 bp in length
21513 21612: gap of 100 bp
21613 22339: contig of 727 bp in length
22340 22439: gap of 100 bp
22440 23119: contig of 680 bp in length
23120 23219: gap of 100 bp
23220 23942: contig of 723 bp in length
23943 24042: gap of 100 bp
24043 24769: contig of 727 bp in length
24770 24869: gap of 100 bp
24870 25551: contig of 682 bp in length
25552 25651: gap of 100 bp
25652 26344: contig of 693 bp in length

* 26345 26444: gap of 100 bp
26445 27141: contig of 697 bp in length
27142 27241: gap of 100 bp
27242 27951: contig of 710 bp in length
27952 28051: gap of 100 bp
28052 28782: contig of 731 bp in length
28783 28882: gap of 100 bp
28883 29600: contig of 718 bp in length
29601 29700: gap of 100 bp
29701 30422: contig of 722 bp in length
30423 30522: gap of 100 bp
30523 31204: contig of 682 bp in length
31205 31304: gap of 100 bp
31305 31989: contig of 685 bp in length
31990 32089: gap of 100 bp
32090 32812: contig of 723 bp in length
32813 32912: gap of 100 bp
32913 33593: contig of 681 bp in length
33594 33693: gap of 100 bp
33694 34407: contig of 714 bp in length
34408 34507: gap of 100 bp
34508 35228: contig of 721 bp in length
35229 35328: gap of 100 bp
35329 36066: contig of 738 bp in length
36067 36166: gap of 100 bp
36167 36850: contig of 684 bp in length
36851 36950: gap of 100 bp
36951 37656: contig of 706 bp in length
37657 37756: gap of 100 bp
37757 38405: contig of 649 bp in length
38406 38505: gap of 100 bp
38506 39193: contig of 688 bp in length
39194 39293: gap of 100 bp
39294 39990: contig of 697 bp in length
39991 40090: gap of 100 bp
40091 40798: contig of 708 bp in length
40799 40898: gap of 100 bp
40899 41582: contig of 684 bp in length
41583 41682: gap of 100 bp
41683 42419: contig of 737 bp in length
42420 42519: gap of 100 bp
42520 43252: contig of 733 bp in length
43253 43352: gap of 100 bp
43353 44024: contig of 672 bp in length
44025 44124: gap of 100 bp
44125 44825: contig of 701 bp in length
44826 44925: gap of 100 bp
44926 45640: contig of 715 bp in length
45641 45740: gap of 100 bp
45741 46438: contig of 698 bp in length
46439 46538: gap of 100 bp
46539 47266: contig of 728 bp in length.

FEATURES
source

1..47266
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP24-90F19"
/clone_lib="RPCI-24 Male Mouse BAC"

BASE COUNT 10916 a 9445 c 9380 g 11605 t 5920 others
ORIGIN

Query Match 100.0%; Score 231; DB 2; Length 47266;
Best Local Similarity 100.0%; Pred. No. 3.1e-65;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGGGTGCGAGCTAAGCAGTGCACAGAGGATAGTCTGTGGCATCTCTGGAAGG 60
|||||

Db 41188 GTGGGTGCGAGCTAAGCAGTGCACAGAGGATAGTCTGTGGCATCTCTGGAAGG 41129
|||||

Qy 61 CAGCTCTGATTGATGAAGTTTCAGTGTCTCTCCCTCCAGGAAACGACACG 120
|||||

Db 41128 CAGCTCTGATTGATGAAGTTTCAGTGTCTCTCCCTCCAGGAAACGACACG 41069
|||||

OY 121 GGAGCTGCCAGACCTCTCTGGGTGATGAGCCCTAAGGATGTTTGTGTAGGGCCCT 180
 Db 41068 GGAGCTGCCAGACCTCTCTGGGTGATGAGCCCTAAGGATGTTTGTGTAGGGCCCT 41009
 OY 181 ATGCTTGACACTGGGATCAGACCTCTACCTTCACCCATGAGCTGCTT 231
 Db 41008 ATGCTTGACACTGGGATCAGACCTCTACCTTCACCCATGAGCTGCTT 40958

RESULT 6
 AC116705 59079 bp DNA linear HTG 16-MAY-2002
 LOCUS Mus musculus clone RP23-26P10, LOW-PASS SEQUENCE SAMPLING.
 DEFINITION AC116705
 ACCESSION AC116705.2 GI:20806391
 VERSION HTG: HTGS_PHASE0.
 KEYWORDS Mus musculus.
 SOURCE Mus musculus
 ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

1 (bases 1 to 59079)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (02-APR-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 59079)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavskiy,L.,
 Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
 Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
 Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
 Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
 JOURNAL
 COMMENT

Direct Submission
 Submitted (16-MAY-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 16, 2002 this sequence version replaced gi:19881948.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L22303
 Center clone name: 26_P_10

 * NOTE: This record contains 73 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.
 * 1 704: contig of 704 bp in length
 * 705 804: gap of 100 bp
 * 805 1517: contig of 713 bp in length
 * 1518 1617: gap of 100 bp
 * 1618 2333: contig of 716 bp in length
 * 2334 2433: gap of 100 bp
 * 2434 3147: contig of 714 bp in length
 * 3148 3247: gap of 100 bp
 * 3248 3944: contig of 697 bp in length
 * 3945 4044: gap of 100 bp
 * 4045 4759: contig of 715 bp in length
 * 4760 4859: gap of 100 bp
 * 4860 5582: contig of 723 bp in length
 * 5583 5682: gap of 100 bp
 * 5683 6374: contig of 692 bp in length
 * 6375 6474: gap of 100 bp
 * 6475 7189: contig of 715 bp in length
 * 7190 7289: gap of 100 bp
 * 7290 7997: contig of 708 bp in length
 * 7998 8097: gap of 100 bp
 * 8098 8817: contig of 720 bp in length
 * 8818 8917: gap of 100 bp
 * 8918 9628: contig of 711 bp in length
 * 9629 9728: gap of 100 bp
 * 9729 10449: contig of 721 bp in length
 * 10450 10549: gap of 100 bp
 * 10550 11284: contig of 735 bp in length
 * 11285 11384: gap of 100 bp
 * 11385 12074: contig of 690 bp in length
 * 12075 12174: gap of 100 bp
 * 12175 12861: contig of 687 bp in length
 * 12862 12961: gap of 100 bp
 * 12962 13627: contig of 666 bp in length
 * 13628 13727: gap of 100 bp
 * 13728 14421: contig of 694 bp in length
 * 14422 14521: gap of 100 bp
 * 14522 15246: contig of 725 bp in length
 * 15247 15346: gap of 100 bp
 * 15347 16075: contig of 729 bp in length
 * 16076 16175: gap of 100 bp
 * 16176 16905: contig of 730 bp in length
 * 16906 17005: gap of 100 bp
 * 17006 17707: contig of 702 bp in length
 * 17708 17807: gap of 100 bp
 * 17808 18510: contig of 703 bp in length
 * 18511 18610: gap of 100 bp

```
* 18611 19327: contig of 717 bp in length
* 19328 19427: gap of 100 bp
* 19428 20107: contig of 680 bp in length
* 20108 20207: gap of 100 bp
* 20208 20925: contig of 718 bp in length
* 20926 21025: gap of 100 bp
* 21026 21738: contig of 713 bp in length
* 21739 21838: gap of 100 bp
* 21839 22538: contig of 700 bp in length
* 22539 22638: gap of 100 bp
* 22639 23357: contig of 719 bp in length
* 23358 23457: gap of 100 bp
* 23458 24172: contig of 715 bp in length
* 24173 24272: gap of 100 bp
* 24273 24966: contig of 694 bp in length
* 24967 25066: gap of 100 bp
* 25067 25770: contig of 704 bp in length
* 25771 25870: gap of 100 bp
* 25871 26564: contig of 694 bp in length
* 26565 26664: gap of 100 bp
* 26665 27385: contig of 721 bp in length
* 27386 27485: gap of 100 bp
* 27486 28191: contig of 706 bp in length
* 28192 28291: gap of 100 bp
* 28292 29005: contig of 714 bp in length
* 29006 29105: gap of 100 bp
* 29106 29813: contig of 708 bp in length
* 29814 29913: gap of 100 bp
* 29914 30637: contig of 724 bp in length
* 30638 30737: gap of 100 bp
* 30738 31450: contig of 713 bp in length
* 31451 31550: gap of 100 bp
* 31551 32267: contig of 717 bp in length
* 32268 32367: gap of 100 bp
* 32368 33077: contig of 710 bp in length
* 33078 33177: gap of 100 bp
* 33178 33888: contig of 711 bp in length
* 33889 33988: gap of 100 bp
* 33989 34706: contig of 718 bp in length
* 34707 34806: gap of 100 bp
* 34807 35494: contig of 688 bp in length
* 35495 35594: gap of 100 bp
* 35595 36298: contig of 704 bp in length
* 36299 36398: gap of 100 bp
* 36399 37108: contig of 710 bp in length
* 37109 37208: gap of 100 bp
* 37209 37899: contig of 691 bp in length
* 37900 37999: gap of 100 bp
* 38000 38685: contig of 686 bp in length
* 38686 38785: gap of 100 bp
* 38786 39514: contig of 729 bp in length
* 39515 39614: gap of 100 bp
* 39615 40321: contig of 707 bp in length
* 40322 40421: gap of 100 bp
* 40422 41135: contig of 714 bp in length
* 41136 41235: gap of 100 bp
* 41236 41943: contig of 708 bp in length
* 41944 42043: gap of 100 bp
* 42044 42770: contig of 727 bp in length
* 42771 42870: gap of 100 bp
* 42871 43602: contig of 732 bp in length
* 43603 43702: gap of 100 bp
* 43703 44419: contig of 717 bp in length
* 44420 44519: gap of 100 bp
* 44520 45230: contig of 711 bp in length
```

Query Match 100.0%; Score 231; DB 2: Length 59079;
Best Local Similarity 100.0%; Pred. No. 3,1e-65;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GTGGGGTCCGAGGCTAGACACTGCACAGAGATAGCTGTGTGGCATCTCTGTGGAGG 60
|||||
DB 34216 GTGGGGTCCGAGGCTAGACACTGCACAGAGATAGCTGTGTGGCATCTCTGTGGAGG 34275
```

```
QY 61 CACCTCTGATTGCATGAAGTTTCAGTGTCTAGTTCCTCTCCCTCCCGAGAAAAGACACG 120
|||||
DB 34276 CACGCTCTGATTGCATGAAGTTTCAGTGTCTAGTTCCTCTCCCTCCCGAGAAAAGACACG 34335
|||||
QY 121 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGCCCCCT 180
|||||
DB 34336 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGCCCCCT 34395
|||||
QY 181 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCATGAGGCTTGCTT 231
|||||
DB 34396 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCATGAGGCTTGCTT 34446
|||||
```

RESULT 7

AC124106/c

LOCUS

DEFINITION

AC124106

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

60588 bp DNA linear HTG 09-JUN-2002
Mus musculus clone RP24-299A7, LOW-PASS SEQUENCE SAMPLING.
AC124106
Mus musculus
AC124106.1 GI:21360210
HTG; HTGS_PHASE0.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 60588)
Mus musculus, clone RP24-299A7
Unpublished
2 (bases 1 to 60588)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhgalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
Johnson, R., Jones, C., Katat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N.,
Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,
Meneus, L., Mihova, T., Mienga, V., Murphy, J., Naylor, J., Nguyen, C.,
Nicot, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (09-JUN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26631
Center clone name: 299_A_7

* NOTE: This record contains 74 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 738: contig of 738 bp in length
739 838: gap of 100 bp
839 1532: contig of 694 bp in length
1533 1632: gap of 100 bp
1633 2363: contig of 731 bp in length
2364 2463: gap of 100 bp
2464 3214: contig of 751 bp in length
3215 3314: gap of 100 bp
3315 4044: contig of 730 bp in length
4045 4144: gap of 100 bp
4145 4874: contig of 730 bp in length
4875 4974: gap of 100 bp
4975 5669: contig of 695 bp in length
5670 5769: gap of 100 bp
5770 6483: contig of 714 bp in length
6484 6583: gap of 100 bp
6584 7310: contig of 727 bp in length
7311 7410: gap of 100 bp
7411 8124: contig of 714 bp in length
8125 8224: gap of 100 bp
8225 8960: contig of 736 bp in length
8961 9060: gap of 100 bp
9061 9792: contig of 732 bp in length
9793 9892: gap of 100 bp
9893 10617: contig of 725 bp in length
10618 10717: gap of 100 bp
10718 11453: contig of 736 bp in length
11454 11553: gap of 100 bp
11554 12274: contig of 721 bp in length
12275 12374: gap of 100 bp
12375 13095: contig of 721 bp in length
13096 13195: gap of 100 bp
13196 13906: contig of 711 bp in length
13907 14006: gap of 100 bp
14007 14709: contig of 703 bp in length
14710 14809: gap of 100 bp
14810 15525: contig of 716 bp in length
15526 15625: gap of 100 bp
15626 16354: contig of 729 bp in length
16355 16454: gap of 100 bp
16455 17188: contig of 734 bp in length
17189 17288: gap of 100 bp
17289 18008: contig of 720 bp in length
18009 18108: gap of 100 bp
18109 18838: contig of 730 bp in length
18839 18938: gap of 100 bp
18939 19677: contig of 739 bp in length
19678 19777: gap of 100 bp
19778 20527: contig of 750 bp in length
20528 20627: gap of 100 bp
20628 21361: contig of 734 bp in length
21362 21461: gap of 100 bp
21462 22147: contig of 686 bp in length
22148 22247: gap of 100 bp
22248 22945: contig of 698 bp in length
22946 23045: gap of 100 bp
23046 23739: contig of 694 bp in length
23740 23839: gap of 100 bp
23840 24556: contig of 717 bp in length
24557 24656: gap of 100 bp
24657 25379: contig of 723 bp in length
25380 25479: gap of 100 bp
25480 26213: contig of 734 bp in length
26214 26313: gap of 100 bp
26314 27037: contig of 724 bp in length
27038 27137: gap of 100 bp
27138 27845: contig of 708 bp in length
27846 27945: gap of 100 bp

27946 28683: contig of 738 bp in length
28684 28783: gap of 100 bp
28784 29513: contig of 730 bp in length
29514 29613: gap of 100 bp
29614 30323: contig of 710 bp in length
30324 30423: gap of 100 bp
30424 31145: contig of 722 bp in length
31146 31245: gap of 100 bp
31246 31967: contig of 722 bp in length
31968 32067: gap of 100 bp
32068 32749: contig of 682 bp in length
32750 32849: gap of 100 bp
32850 33570: contig of 721 bp in length
33571 33670: gap of 100 bp
33671 34395: contig of 725 bp in length
34396 34495: gap of 100 bp
34496 35223: contig of 728 bp in length
35224 35323: gap of 100 bp
35324 36039: contig of 716 bp in length
36040 36139: gap of 100 bp
36140 36864: contig of 725 bp in length
36865 36964: gap of 100 bp
36965 37702: contig of 738 bp in length
37703 37802: gap of 100 bp
37803 38536: contig of 734 bp in length
38537 38636: gap of 100 bp
38637 39332: contig of 696 bp in length
39333 39432: gap of 100 bp
39433 40128: contig of 696 bp in length
40129 40228: gap of 100 bp
40229 40950: contig of 722 bp in length
40951 41050: gap of 100 bp
41051 41772: contig of 722 bp in length
41773 41872: gap of 100 bp
41873 42581: contig of 709 bp in length
42582 42681: gap of 100 bp
42682 43417: contig of 736 bp in length
43418 43517: gap of 100 bp
43518 44245: contig of 728 bp in length
44246 44345: gap of 100 bp
44346 45073: contig of 728 bp in length
45074 45173: gap of 100 bp
45174 45879: contig of 706 bp in length
45880 45979: gap of 100 bp
45980 46709: contig of 730 bp in length
46710 46809: gap of 100 bp
46810 47522: contig of 713 bp in length
47523 47622: gap of 100 bp
47623 48321: contig of 699 bp in length
48322 48421: gap of 100 bp
48422 49146: contig of 725 bp in length
49147 49246: gap of 100 bp
49247 49961: contig of 715 bp in length
49962 50061: gap of 100 bp
50062 50798: contig of 737 bp in length
50799 50898: gap of 100 bp
50899 51631: contig of 733 bp in length
51632 51731: gap of 100 bp
51732 52434: contig of 703 bp in length
52435 52534: gap of 100 bp
52535 53256: contig of 722 bp in length
53257 53356: gap of 100 bp
53357 54097: contig of 741 bp in length
54098 54197: gap of 100 bp
54198 54914: contig of 717 bp in length
54915 55014: gap of 100 bp
55015 55734: contig of 720 bp in length
55735 55834: gap of 100 bp
55835 56549: contig of 715 bp in length
56550 56649: gap of 100 bp

Query Match 100.0%; Score 231; DB 2; Length 60588;
Best Local Similarity 100.0%; Pred. No. 3.1e-65;

Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTCGAGGCTAAGCACTGCACAGAGATAGCTGTGTGGCATCTGTGGAAGG 60
 Db 44018 GTGGGTCGAGGCTAAGCACTGCACAGAGATAGCTGTGTGGCATCTGTGGAAGG 43959

QY 61 CACCTCTGATTGTCATGAAGTTTCAGTCTCCTAGTTCCTCCCGGAGAAAACGACACG 120
 Db 43958 CACCTCTGATTGTCATGAAGTTTCAGTCTCCTAGTTCCTCCCGGAGAAAACGACACG 43899

QY 121 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGGCCCT 180
 Db 43898 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGGTTTGTGTAGGGCCCT 43839

QY 181 ATGCTTGACACTGGGATCAGACCTTACCTTACCCTAGGCTTGCTT 231
 Db 43838 ATGCTTGACACTGGGATCAGACCTTACCTTACCCTAGGCTTGCTT 43788

RESULT 8
 AC113184
 LOCUS
 DEFINITION Mus musculus clone RP24-76E4, LOW-PASS SEQUENCE SAMPLING.
 ACCESSION AC113184
 VERSION AC113184.1 GI:18921362
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 61926)
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.
 Mus musculus, clone RP24-76E4
 Unpublished
 2 (bases 1 to 61926)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
 Landers, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C.,
 Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M.,
 McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T.,
 Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
 Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (26-FEB-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L21171
 Center clone name: 76_E_4

* NOTE: This record contains 78 individual

* sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

1 706: contig of 706 bp in length
 * 707 806: gap of 100 bp
 * 807 1500: contig of 694 bp in length
 * 1501 1600: gap of 100 bp
 * 1601 2307: contig of 707 bp in length
 * 2308 2407: gap of 100 bp
 * 2408 3104: contig of 697 bp in length
 * 3105 3204: gap of 100 bp
 * 3205 3894: contig of 690 bp in length
 * 3895 3994: gap of 100 bp
 * 3995 4659: contig of 665 bp in length
 * 4660 4759: gap of 100 bp
 * 4760 5450: contig of 691 bp in length
 * 5451 5550: gap of 100 bp
 * 5551 6260: contig of 710 bp in length
 * 6261 6360: gap of 100 bp
 * 6361 7061: contig of 701 bp in length
 * 7062 7161: gap of 100 bp
 * 7162 7862: contig of 701 bp in length
 * 7863 7962: gap of 100 bp
 * 7963 8648: contig of 686 bp in length
 * 8649 8748: gap of 100 bp
 * 8749 9447: contig of 699 bp in length
 * 9448 9547: gap of 100 bp
 * 9548 10239: contig of 692 bp in length
 * 10240 10339: gap of 100 bp
 * 10340 11042: contig of 703 bp in length
 * 11043 11142: gap of 100 bp
 * 11143 11833: contig of 691 bp in length
 * 11834 11933: gap of 100 bp
 * 11934 12616: contig of 683 bp in length
 * 12617 12716: gap of 100 bp
 * 12717 13426: contig of 710 bp in length
 * 13427 13526: gap of 100 bp
 * 13527 14229: contig of 703 bp in length
 * 14230 14329: gap of 100 bp
 * 14330 15007: contig of 678 bp in length
 * 15008 15107: gap of 100 bp
 * 15108 15799: contig of 692 bp in length
 * 15800 15899: gap of 100 bp
 * 15900 16590: contig of 691 bp in length
 * 16591 16690: gap of 100 bp
 * 16691 17383: contig of 693 bp in length
 * 17384 17483: gap of 100 bp
 * 17484 18197: contig of 714 bp in length
 * 18198 18297: gap of 100 bp
 * 18298 18993: contig of 696 bp in length
 * 18994 19093: gap of 100 bp
 * 19094 19789: contig of 696 bp in length
 * 19790 19889: gap of 100 bp
 * 19890 20583: contig of 694 bp in length
 * 20584 20683: gap of 100 bp
 * 20684 21364: contig of 681 bp in length
 * 21365 21464: gap of 100 bp
 * 21465 22176: contig of 712 bp in length
 * 22177 22276: gap of 100 bp
 * 22277 22973: contig of 697 bp in length
 * 22974 23073: gap of 100 bp
 * 23074 23754: contig of 681 bp in length
 * 23755 23854: gap of 100 bp
 * 23855 24549: contig of 695 bp in length
 * 24550 24649: gap of 100 bp
 * 24650 25368: contig of 719 bp in length

* 25369 25468: gap of 100 bp in length
 * 25469 26170: contig of 702 bp in length
 * 26171 26270: gap of 100 bp
 * 26271 26962: contig of 692 bp in length
 * 26963 27062: gap of 100 bp
 * 27063 27752: contig of 690 bp in length
 * 27753 27852: gap of 100 bp
 * 27853 28550: contig of 698 bp in length
 * 28551 28650: gap of 100 bp
 * 28651 29366: contig of 716 bp in length
 * 29367 29466: gap of 100 bp
 * 29467 30162: contig of 696 bp in length
 * 30163 30262: gap of 100 bp in length
 * 30263 30966: contig of 704 bp in length
 * 30967 31066: gap of 100 bp
 * 31067 31753: contig of 687 bp in length
 * 31754 31853: gap of 100 bp
 * 31854 32555: contig of 702 bp in length
 * 32556 32655: gap of 100 bp
 * 32656 33347: contig of 692 bp in length
 * 33348 33447: gap of 100 bp
 * 33448 34149: contig of 702 bp in length
 * 34150 34249: gap of 100 bp
 * 34250 34957: contig of 708 bp in length
 * 34958 35057: gap of 100 bp
 * 35058 35758: contig of 701 bp in length
 * 35759 35858: gap of 100 bp
 * 35859 36535: contig of 677 bp in length
 * 36536 36635: gap of 100 bp
 * 36636 37331: contig of 696 bp in length
 * 37332 37431: gap of 100 bp
 * 37432 38129: contig of 698 bp in length
 * 38130 38229: gap of 100 bp
 * 38230 38939: contig of 710 bp in length
 * 38940 39039: gap of 100 bp
 * 39040 39733: contig of 694 bp in length
 * 39734 39833: gap of 100 bp
 * 39834 40530: contig of 697 bp in length
 * 40531 40630: gap of 100 bp
 * 40631 41283: contig of 653 bp in length
 * 41284 41383: gap of 100 bp
 * 41384 42090: contig of 707 bp in length
 * 42091 42190: gap of 100 bp
 * 42191 42905: contig of 715 bp in length
 * 42906 43005: gap of 100 bp
 * 43006 43710: contig of 705 bp in length
 * 43711 43810: gap of 100 bp
 * 43811 44502: contig of 692 bp in length
 * 44503 44602: gap of 100 bp
 * 44603 45294: contig of 692 bp in length
 * 45295 45394: gap of 100 bp
 * 45395 46110: contig of 716 bp in length
 * 46111 46210: gap of 100 bp
 * 46211 46922: contig of 712 bp in length
 * 46923 47022: gap of 100 bp
 * 47023 47707: contig of 685 bp in length
 * 47708 47807: gap of 100 bp
 * 47808 48512: contig of 705 bp in length
 * 48513 48612: gap of 100 bp
 * 48613 49295: contig of 683 bp in length
 * 49296 49395: gap of 100 bp
 * 49396 50097: contig of 702 bp in length
 * 50098 50197: gap of 100 bp
 * 50198 50896: contig of 699 bp in length
 * 50897 50996: gap of 100 bp
 * 50997 51679: contig of 683 bp in length
 * 51680 51779: gap of 100 bp
 * 51780 52472: contig of 693 bp in length
 * 52473 52572: gap of 100 bp
 * 52573 53257: contig of 685 bp in length
 * 53258 53357: gap of 100 bp
 * 53358 54057: contig of 700 bp in length
 * 54058 54157: gap of 100 bp

* 54158 54869: contig of 712 bp in length
 * 54870 54969: gap of 100 bp

Query Match 100.0%; Score 231; DB 2; Length 61926;
 Best Local Similarity 100.0%; Pred. No. 31e-65;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GTGGGGTCGAGGCTAAGCACTGCACAGAGATAGCTTGTGGCATCTGTGGAAGG 60
 Db 2695 GTGGGGTCGAGGCTAAGCACTGCACAGAGATAGCTTGTGGCATCTGTGGAAGG 2754
 Oy 61 CACGTCCTGATTGCATGAAGGTTCACTGTCCTAGTTCCTTCCCCCAGAAAACACACAG 120
 Db 2755 CACGTCCTGATTGCATGAAGGTTCACTGTCCTAGTTCCTTCCCCCAGAAAACACACAG 2814
 Oy 121 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTAAGGATGGTGTGTAGGGCCCT 180
 Db 2815 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTAAGGATGGTGTGTAGGGCCCT 2874
 Oy 181 ATGCTTGACACTGGGGATCAGACCTCTACCTTACCCATGAGGCTTCTT 231
 Db 2875 ATGCTTGACACTGGGGATCAGACCTCTACCTTACCCATGAGGCTTCTT 2925

RESULT 9
 AC101394 64042 bp DNA linear HTG 23-NOV-2001
 LOCUS Mus musculus clone RP23-119F14, LOW-PASS SEQUENCE SAMPLING.
 DEFINITION AC101394
 ACCESSION AC101394
 VERSION AC101394.1 GI:17060169
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Mus musculus
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP23-119F14
 Unpublished
 2 (bases 1 to 64042)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
 Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kamat,A., Karatas,A., Kellis,C., LaRoque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Melárim,J.,
 Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Triggilio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu

```
----- Project Information
Center project name: L16611
Center clone name: 119 F 14
```

* NOTE: This record contains 79 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

693	792	692:	contig	of	692	bp	in	length
793	1511:	contig	of	100	bp			
1512	1611:	gap	of	100	bp			
1612	2313:	contig	of	702	bp	in	length	
2314	2413:	gap	of	100	bp			
2414	3108:	contig	of	695	bp	in	length	
3109	3208:	gap	of	100	bp			
3209	3529:	contig	of	721	bp	in	length	
3930	4029:	gap	of	100	bp			
4030	4745:	contig	of	716	bp	in	length	
4746	4845:	gap	of	100	bp			
4846	5559:	contig	of	714	bp	in	length	
5560	5599:	gap	of	100	bp			
5660	6404:	contig	of	745	bp	in	length	
6405	6504:	gap	of	100	bp			
6505	7235:	contig	of	731	bp	in	length	
7236	7335:	gap	of	100	bp			
7336	8068:	contig	of	733	bp	in	length	
8069	8168:	gap	of	100	bp			
8169	8875:	contig	of	707	bp	in	length	
8876	8975:	gap	of	100	bp			
8976	9697:	contig	of	722	bp	in	length	
9698	9797:	gap	of	100	bp			
9798	10508:	contig	of	711	bp	in	length	
10509	10608:	gap	of	100	bp			
10609	11313:	contig	of	705	bp	in	length	
11314	11413:	gap	of	100	bp			
11414	12111:	contig	of	698	bp	in	length	
12112	12211:	gap	of	100	bp			
12212	12526:	contig	of	715	bp	in	length	
12927	13026:	gap	of	100	bp			
13027	13733:	contig	of	707	bp	in	length	
13734	13833:	gap	of	100	bp			
13834	14544:	contig	of	711	bp	in	length	
14545	14644:	gap	of	100	bp			
14645	15345:	contig	of	701	bp	in	length	
15346	15445:	gap	of	100	bp			
15446	16168:	contig	of	723	bp	in	length	
16169	16268:	gap	of	100	bp			
16269	16970:	contig	of	702	bp	in	length	
16971	17070:	gap	of	100	bp			
17071	17764:	contig	of	694	bp	in	length	
17765	17864:	gap	of	100	bp			
17865	18570:	contig	of	706	bp	in	length	
18571	18670:	gap	of	100	bp			
18671	19389:	contig	of	719	bp	in	length	
19390	19489:	gap	of	100	bp			
19490	20197:	contig	of	708	bp	in	length	
20198	20297:	gap	of	100	bp			
20298	21030:	contig	of	733	bp	in	length	
21031	21130:	gap	of	100	bp			
21131	21859:	contig	of	729	bp	in	length	
21860	21959:	gap	of	100	bp			
21960	22675:	contig	of	716	bp	in	length	
22676	22775:	gap	of	100	bp			
22776	23496:	contig	of	721	bp	in	length	
23497	23596:	gap	of	100	bp			

23597	24307:	contig of 711 bp	in length
24308	24407:	gap of 100 bp	
24408	25127:	contig of 720 bp	in length
25128	25227:	gap of 100 bp	
25228	25941:	contig of 714 bp	in length
25942	26041:	gap of 100 bp	
26042	26758:	contig of 717 bp	in length
26759	26858:	gap of 100 bp	
26859	27583:	contig of 725 bp	in length
27584	27683:	gap of 100 bp	
27684	28395:	contig of 712 bp	in length
28396	28495:	gap of 100 bp	
28496	29167:	contig of 672 bp	in length
29168	29267:	gap of 100 bp	
29268	29918:	contig of 651 bp	in length
29919	30018:	gap of 100 bp	
30019	30734:	contig of 716 bp	in length
30735	30834:	gap of 100 bp	
30835	31573:	contig of 738 bp	in length
31573	31672:	gap of 100 bp	
31673	32377:	contig of 705 bp	in length
32378	32477:	gap of 100 bp	
32478	33193:	contig of 716 bp	in length
33194	33293:	gap of 100 bp	
33294	33982:	contig of 689 bp	in length
33983	34082:	gap of 100 bp	
34083	34795:	contig of 713 bp	in length
34796	34895:	gap of 100 bp	
34896	35603:	contig of 708 bp	in length
35604	35703:	gap of 100 bp	
35704	36427:	contig of 724 bp	in length
36428	36527:	gap of 100 bp	
36528	37256:	contig of 729 bp	in length
37257	37356:	gap of 100 bp	
37357	38073:	contig of 723 bp	in length
38080	38179:	gap of 100 bp	
38180	38900:	contig of 721 bp	in length
38901	39000:	gap of 100 bp	
39001	39717:	contig of 717 bp	in length
39718	39817:	gap of 100 bp	
39818	40532:	contig of 715 bp	in length
40533	40632:	gap of 100 bp	
40633	41342:	contig of 710 bp	in length
41343	41442:	gap of 100 bp	
41443	42138:	contig of 696 bp	in length
42139	42238:	gap of 100 bp	
42239	42942:	contig of 704 bp	in length
42943	43042:	gap of 100 bp	
43043	43741:	contig of 699 bp	in length
43742	43841:	gap of 100 bp	
43842	44559:	contig of 718 bp	in length
44560	44659:	gap of 100 bp	
44660	45353:	contig of 694 bp	in length
45354	45453:	gap of 100 bp	
45454	46174:	contig of 721 bp	in length
46175	46274:	gap of 100 bp	
46275	46989:	contig of 715 bp	in length
46990	47089:	gap of 100 bp	
47090	47789:	contig of 700 bp	in length
47790	47889:	gap of 100 bp	
47890	48602:	contig of 713 bp	in length
48603	48702:	gap of 100 bp	
48703	49409:	contig of 707 bp	in length
49410	49509:	gap of 100 bp	
49510	50188:	contig of 679 bp	in length
50189	50288:	gap of 100 bp	
50289	51005:	contig of 717 bp	in length
51006	51105:	gap of 100 bp	
51106	51823:	contig of 720 bp	in length
51826	51925:	gap of 100 bp	
51926	52629:	contig of 704 bp	in length
52630	52729:	gap of 100 bp	
52730	53448:	contig of 719 bp	in length

* 53449 53548: gap of 100 bp
* 53549 54253: contig of 711 bp in length
* 54260 54359: gap of 100 bp
* 54360 55074: contig of 715 bp in length
* 55075 55174: gap of 100 bp
* 55175 55883: contig of 709 bp in length
* 55884 55983: gap of 100 bp

Query Match 100.0%; Score 231; DB 2; Length 64042;
Best Local Similarity 100.0%; Pred. No. 3.1e-65;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGGGTGCGAGGCTAGCACTGCACAGATAGCTGTGTTGGCATCTCTGGAGG 60
Db 32664 GTGGGGTGGAGGCTAGCACTGCACAGATAGCTGTGTTGGCATCTCTGGAGG 32723
Qy 61 CACGCTGATTCGATGAAGGTTCAAGTTCCTTCCCTCCCGAGGAAACGACACG 120
Db 32724 CACGCTGATTCGATGAAGGTTCAAGTTCCTTCCCTCCCGAGGAAACGACACG 32783
Qy 121 GGAGTGGCCCAAGACCTCTCTGGGTGATGAGCCCTTAAGGATGTTGTGTAGGCCCTT 180
Db 32784 GGAGTGGCCCAAGACCTCTCTGGGTGATGAGCCCTTAAGGATGTTGTGTAGGCCCTT 32843
Qy 181 ATGCTGCACACTGGGATCAGACTCTACCTTCACCATGAGCTTCTT 231
Db 32844 ATGCTGCACACTGGGATCAGACTCTACCTTCACCATGAGCTTCTT 32894

RESULT 10
AC100424/c

LOCUS AC100424 66393 bp DNA linear HTG 22-NOV-2001
DEFINITION Mus musculus clone RP23-136F8, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC100424
VERSION AC100424.1 GI:17047790
KEYWORDS HTG; HTGS, PHASED.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Blurren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-136F8
Unpublished
2 (bases 1 to 66393)
Blurren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferrelira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Mencus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,K., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

TITLE
JOURNAL
COMMENT
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15312
Center clone name: 136_F_8

* NOTE: This record contains 84 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 692: contig of 692 bp in length
* 693 792: gap of 100 bp
* 793 1509: contig of 717 bp in length
* 1510 1609: gap of 100 bp
* 1610 2315: contig of 706 bp in length
* 2316 2415: gap of 100 bp
* 2416 3132: contig of 717 bp in length
* 3133 3232: gap of 100 bp
* 3233 3912: contig of 680 bp in length
* 3913 4012: gap of 100 bp
* 4013 4682: contig of 870 bp in length
* 4683 4782: gap of 100 bp
* 4783 5477: contig of 695 bp in length
* 5478 5577: gap of 100 bp
* 5578 6280: contig of 703 bp in length
* 6281 6380: gap of 100 bp
* 6381 7086: contig of 886 bp in length
* 7087 7166: gap of 100 bp
* 7167 7870: contig of 704 bp in length
* 7871 7970: gap of 100 bp
* 7971 8665: contig of 695 bp in length
* 8666 8765: gap of 100 bp
* 8766 9472: contig of 707 bp in length
* 9473 9572: gap of 100 bp
* 9573 10230: contig of 658 bp in length
* 10231 10330: gap of 100 bp
* 10331 11024: contig of 694 bp in length
* 11025 11124: gap of 100 bp
* 11125 11809: contig of 885 bp in length
* 11810 11909: gap of 100 bp
* 11910 12589: contig of 680 bp in length
* 12590 12689: gap of 100 bp
* 12690 13374: contig of 685 bp in length
* 13375 13474: gap of 100 bp
* 13475 14170: contig of 896 bp in length
* 14171 14270: gap of 100 bp
* 14271 14967: contig of 697 bp in length
* 14968 15067: gap of 100 bp
* 15068 15770: contig of 703 bp in length
* 15771 15870: gap of 100 bp
* 15871 16588: contig of 718 bp in length
* 16589 16688: gap of 100 bp
* 16689 17357: contig of 669 bp in length
* 17358 17457: gap of 100 bp
* 17458 18160: contig of 703 bp in length
* 18161 18260: gap of 100 bp
* 18261 18972: contig of 712 bp in length
* 18973 19072: gap of 100 bp
* 19073 19724: contig of 652 bp in length
* 19725 19824: gap of 100 bp
* 19825 20520: contig of 696 bp in length
* 20521 20620: gap of 100 bp
* 20621 21326: contig of 706 bp in length

* 21327 21426: gap of 100 bp
* 21427 22117: contig of 691 bp in length
* 22118 22217: gap of 100 bp
* 22218 22904: contig of 687 bp in length
* 22905 23004: gap of 100 bp
* 23005 23705: contig of 701 bp in length
* 23706 23805: gap of 100 bp
* 23806 24497: contig of 692 bp in length
* 24498 24597: gap of 100 bp
* 24598 25272: contig of 675 bp in length
* 25273 25372: gap of 100 bp
* 25373 26083: contig of 711 bp in length
* 26084 26183: gap of 100 bp
* 26184 26860: contig of 677 bp in length
* 26861 26960: gap of 100 bp
* 26961 27628: contig of 668 bp in length
* 27629 27728: gap of 100 bp
* 27729 28417: contig of 689 bp in length
* 28418 28517: gap of 100 bp
* 28518 29199: contig of 682 bp in length
* 29200 29299: gap of 100 bp
* 29300 29957: contig of 658 bp in length
* 29958 30057: gap of 100 bp
* 30058 30742: contig of 685 bp in length
* 30743 30842: gap of 100 bp
* 30843 31537: contig of 695 bp in length
* 31538 31637: gap of 100 bp
* 31638 32348: contig of 711 bp in length
* 32349 32448: gap of 100 bp
* 32449 33128: contig of 680 bp in length
* 33129 33228: gap of 100 bp
* 33229 33910: contig of 682 bp in length
* 33911 34010: gap of 100 bp
* 34011 34723: contig of 713 bp in length
* 34724 34823: gap of 100 bp
* 34824 35524: contig of 701 bp in length
* 35525 35624: gap of 100 bp
* 35625 36304: contig of 680 bp in length
* 36305 36404: gap of 100 bp
* 36405 37092: contig of 688 bp in length
* 37093 37192: gap of 100 bp
* 37193 37879: contig of 687 bp in length
* 37880 37979: gap of 100 bp
* 37980 38665: contig of 686 bp in length
* 38666 38765: gap of 100 bp
* 38766 39434: contig of 669 bp in length
* 39435 39534: gap of 100 bp
* 39535 40251: contig of 717 bp in length
* 40252 40351: gap of 100 bp
* 40352 41047: contig of 696 bp in length
* 41048 41147: gap of 100 bp
* 41148 41821: contig of 674 bp in length
* 41822 41921: gap of 100 bp
* 41922 42631: contig of 710 bp in length
* 42632 42731: gap of 100 bp
* 42732 43437: contig of 706 bp in length
* 43438 43537: gap of 100 bp
* 43538 44243: contig of 706 bp in length
* 44244 44343: gap of 100 bp
* 44344 45011: contig of 668 bp in length
* 45012 45111: gap of 100 bp
* 45112 45789: contig of 678 bp in length
* 45790 45889: gap of 100 bp
* 45890 46583: contig of 694 bp in length
* 46584 46683: gap of 100 bp
* 46684 47363: contig of 680 bp in length
* 47364 47463: gap of 100 bp
* 47464 48166: contig of 703 bp in length
* 48167 48266: gap of 100 bp
* 48267 48971: contig of 705 bp in length
* 48972 49071: gap of 100 bp
* 49072 49774: contig of 703 bp in length
* 49775 49874: gap of 100 bp

* 49875 50562: contig of 688 bp in length
* 50563 50662: gap of 100 bp
* 50663 51336: contig of 674 bp in length
* 51337 51436: gap of 100 bp
* 51437 52136: contig of 700 bp in length
* 52137 52236: gap of 100 bp
* 52237 52967: contig of 731 bp in length
* 52968 53067: gap of 100 bp
* 53068 53748: contig of 681 bp in length
* 53749 53848: gap of 100 bp
* 53849 54539: contig of 691 bp in length
* 54540 54639: gap of 100 bp

Query Match 100.0%; Score 231; DB 2; Length 66393;
Best Local Similarity 100.0%; Pred. No. 3,le-65;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGCGAGGCTAAGCAGCTGACAGAGGATAGCTTGGTGGGATCCTCTGGGAAGG 60
|||||
DB 22021 GTGGGTGCGAGGCTAAGCAGCTGACAGAGGATAGCTTGGTGGGATCCTCTGGGAAGG 21962
QY 61 CAGCTCTGATTGATGATGAAGGTTTCAGTGTCTCTAGTTCCTTCCCTCCAGGAAAAGACGACG 120
|||||
DB 21961 CAGCTCTGATTGATGATGAAGGTTTCAGTGTCTCTAGTTCCTTCCCTCCAGGAAAAGACGACG 21902
QY 121 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCCTAAGGATGGTTTGTGTAGGGCCCT 180
|||||
DB 21901 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCCTAAGGATGGTTTGTGTAGGGCCCT 21842
QY 181 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCCACCATGAGGCTTGCCTT 231
|||||
DB 21841 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCCACCATGAGGCTTGCCTT 21791

RESULT 11
AC101278
LOCUS
DEFINITION Mus musculus clone RP23-101P16, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC101278
VERSION AC101278.1 GI:17060053
KEYWORDS HTG: HTGS_PHASE0.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 66489)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-101P16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 66489)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., Labouque,K.,
Lamazares,R., Landers,T., Lehotzky,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Olliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission

TITLE

JOURNAL
COMMENT

Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www.seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L16354
Center clone name: 101_P_16

* NOTE: This record contains 81 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 723: contig of 723 bp in length
* 724 823: gap of 100 bp
* 824 1542: contig of 719 bp in length
* 1543 1642: gap of 100 bp
* 1643 2341: contig of 699 bp in length
* 2342 2441: gap of 100 bp
* 2442 3174: contig of 733 bp in length
* 3175 3274: gap of 100 bp
* 3275 4014: contig of 740 bp in length
* 4015 4114: gap of 100 bp
* 4115 4846: contig of 732 bp in length
* 4847 4946: gap of 100 bp
* 4947 5656: contig of 710 bp in length
* 5657 5756: gap of 100 bp
* 5757 6473: contig of 717 bp in length
* 6474 6573: gap of 100 bp
* 6574 7317: contig of 744 bp in length
* 7318 7417: gap of 100 bp
* 7418 8129: contig of 712 bp in length
* 8130 8229: gap of 100 bp
* 8230 8953: contig of 724 bp in length
* 8954 9053: gap of 100 bp
* 9054 9838: contig of 785 bp in length
* 9839 9938: gap of 100 bp
* 9939 10667: contig of 729 bp in length
* 10668 10767: gap of 100 bp
* 10768 11503: contig of 736 bp in length
* 11504 11603: gap of 100 bp
* 11604 12346: contig of 743 bp in length
* 12347 12446: gap of 100 bp
* 12447 13168: contig of 722 bp in length
* 13169 13268: gap of 100 bp
* 13269 13995: contig of 727 bp in length
* 13996 14095: gap of 100 bp
* 14096 14810: contig of 715 bp in length
* 14811 14910: gap of 100 bp
* 14911 15587: contig of 677 bp in length
* 15588 15687: gap of 100 bp
* 15688 16390: contig of 703 bp in length
* 16391 16490: gap of 100 bp
* 16491 17212: contig of 722 bp in length
* 17213 17312: gap of 100 bp
* 17313 18033: contig of 721 bp in length
* 18034 18133: gap of 100 bp
* 18134 18856: contig of 723 bp in length
* 18857 18956: gap of 100 bp
* 18957 19693: contig of 737 bp in length
* 19694 19793: gap of 100 bp

* 19794 20529: contig of 736 bp in length
* 20530 20629: gap of 100 bp
* 20630 21364: contig of 735 bp in length
* 21365 21464: gap of 100 bp
* 21465 22189: contig of 725 bp in length
* 22190 22289: gap of 100 bp
* 22290 22981: contig of 692 bp in length
* 22982 23081: gap of 100 bp
* 23082 23780: contig of 699 bp in length
* 23781 23880: gap of 100 bp
* 23881 24606: contig of 726 bp in length
* 24607 24706: gap of 100 bp
* 24707 25424: contig of 718 bp in length
* 25425 25524: gap of 100 bp
* 25525 26243: contig of 719 bp in length
* 26244 26343: gap of 100 bp
* 26344 27068: contig of 725 bp in length
* 27069 27168: gap of 100 bp
* 27169 27900: contig of 732 bp in length
* 27901 28000: gap of 100 bp
* 28001 28726: contig of 726 bp in length
* 28727 28826: gap of 100 bp
* 28827 29559: contig of 733 bp in length
* 29560 29659: gap of 100 bp
* 29660 30397: contig of 738 bp in length
* 30398 30497: gap of 100 bp
* 30498 31212: contig of 715 bp in length
* 31213 31312: gap of 100 bp
* 31313 32025: contig of 713 bp in length
* 32026 32125: gap of 100 bp
* 32126 32855: contig of 730 bp in length
* 32856 32955: gap of 100 bp
* 32956 33674: contig of 719 bp in length
* 33675 33774: gap of 100 bp
* 33775 34480: contig of 706 bp in length
* 34481 34580: gap of 100 bp
* 34581 35294: contig of 714 bp in length
* 35295 35394: gap of 100 bp
* 35395 36121: contig of 727 bp in length
* 36122 36221: gap of 100 bp
* 36222 36952: contig of 731 bp in length
* 36953 37052: gap of 100 bp
* 37053 37771: contig of 719 bp in length
* 37772 37871: gap of 100 bp
* 37872 38601: contig of 730 bp in length
* 38602 38701: gap of 100 bp
* 38702 39431: contig of 730 bp in length
* 39432 39531: gap of 100 bp
* 39532 40256: contig of 725 bp in length
* 40257 40356: gap of 100 bp
* 40357 41067: contig of 711 bp in length
* 41068 41167: gap of 100 bp
* 41168 41897: contig of 730 bp in length
* 41898 41997: gap of 100 bp
* 41998 42712: contig of 715 bp in length
* 42713 42812: gap of 100 bp
* 42813 43533: contig of 721 bp in length
* 43534 43633: gap of 100 bp
* 43634 44354: contig of 721 bp in length
* 44355 44454: gap of 100 bp
* 44455 45187: contig of 733 bp in length
* 45188 45287: gap of 100 bp
* 45288 46026: contig of 739 bp in length
* 46027 46126: gap of 100 bp
* 46127 46824: contig of 698 bp in length
* 46825 46924: gap of 100 bp
* 46925 47656: contig of 732 bp in length
* 47657 47756: gap of 100 bp
* 47757 48484: contig of 728 bp in length
* 48485 48584: gap of 100 bp
* 48585 49301: contig of 717 bp in length
* 49302 49401: gap of 100 bp
* 49402 50119: contig of 718 bp in length

* 50120 50219: gap of 100 bp
 * 50220 50334: contig of 715 bp in length
 * 50935 51034: gap of 100 bp
 * 51035 51766: contig of 732 bp in length
 * 51767 51866: gap of 100 bp
 * 51867 52575: contig of 709 bp in length
 * 52576 52675: gap of 100 bp
 * 52676 53400: contig of 725 bp in length
 * 53401 53500: gap of 100 bp
 * 53501 54217: contig of 717 bp in length
 * 54218 54317: gap of 100 bp
 * 54318 55003: contig of 686 bp in length
 * 55004 55103: gap of 100 bp
 * 55104 55811: contig of 708 bp in length
 * 55812 55911: gap of 100 bp
 * 55912 56634: contig of 723 bp in length
 * 56635 56734: gap of 100 bp

Query Match 100.0% Score 231: DB 2: Length 66489;
 Best Local Similarity 100.0%; Pred. No. 3.le-65;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGGAGGCAAGCACTGCACAGGATAGCTTCTGTGGCATCCTGTGGAAGG 60
 |||||
 Db 19918 GTGGGTGGAGGCAAGCACTGCACAGGATAGCTTCTGTGGCATCCTGTGGAAGG 19977
 |||||

QY 61 CAGCTCTGATTCGATGAAGGTTTCAGTGTCTAGTCCCTTCCCTCCAGGAAACACGACG 120
 |||||
 Db 19978 CAGCTCTGATTCGATGAAGGTTTCAGTGTCTAGTCCCTTCCCTCCAGGAAACACGACG 20037
 |||||

QY 121 GGAGCTGGCCAGAGCTCTCTGGGTGATGAGCTAGGATGTTTGTGTAGGGCCCT 180
 |||||
 Db 20038 GGAGCTGGCCAGAGCTCTCTGGGTGATGAGCTAGGATGTTTGTGTAGGGCCCT 20097
 |||||

QY 181 ATGCTTGACACTGGGATGCACACCTTACCTTACCCATGAGGCTTGCTT 231
 |||||
 Db 20098 ATGCTTGACACTGGGATGCACACCTTACCTTACCCATGAGGCTTGCTT 20148
 |||||

RESULT 12
 AC101278/c
 LOCUS AC101278 66489 bp DNA linear HTG 23-NOV-2001
 DEFINITION Mus musculus clone RP23-101P16, LOW-PASS SEQUENCE SAMPLING.
 ACCESSION AC101278
 VERSION AC101278.1 GI:17060053
 KEYWORDS HTG; HTGS_PHASE0.
 SOURCE Mus musculus.
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 66489)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Mus musculus, clone RP23-101P16
 Unpublished
 2 (bases 1 to 66489)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
 Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
 Brown,A., Camarata,J., Campolano,A., Chang,J., Chazaro,B.,
 Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
 Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
 Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K.,
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
 MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
 McCarthy,M., McEwan,P., McKernan,K., McPheters,R., Meldrim,J.,
 Meneus,L., Mihova,I., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
 Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
 Topham,K., Travers,M., Travis,N., Trigglio,J., Vassiliev,H.,
 Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: LI6354
 Center clone name: 101_P16

* NOTE: This record contains 81 individual
 * sequencing reads that have not been assembled into
 * contigs. Runs of N are used to separate the reads
 * and the order in which they appear is completely
 * arbitrary. Low-pass sequence sampling is useful for
 * identifying clones that may be gene-rich and allows
 * overlap relationships among clones to be deduced.
 * However, it should not be assumed that this clone
 * will be sequenced to completion. In the event that
 * the record is updated, the accession number will
 * be preserved.

* 1 723: contig of 723 bp in length
 * 724 823: gap of 100 bp
 * 824 1542: contig of 719 bp in length
 * 1543 1642: gap of 100 bp
 * 1643 2341: contig of 699 bp in length
 * 2342 2441: gap of 100 bp
 * 2442 3174: contig of 733 bp in length
 * 3175 3274: gap of 100 bp
 * 3275 4014: contig of 740 bp in length
 * 4015 4114: gap of 100 bp
 * 4115 4846: contig of 732 bp in length
 * 4847 4946: gap of 100 bp
 * 4947 5656: contig of 710 bp in length
 * 5657 5756: gap of 100 bp
 * 5757 6473: contig of 717 bp in length
 * 6474 6573: gap of 100 bp
 * 6574 7317: contig of 744 bp in length
 * 7318 7417: gap of 100 bp
 * 7418 8129: contig of 712 bp in length
 * 8130 8229: gap of 100 bp
 * 8230 8953: contig of 724 bp in length
 * 8954 9053: gap of 100 bp
 * 9054 9838: contig of 785 bp in length
 * 9839 9938: gap of 100 bp
 * 9939 10667: contig of 729 bp in length
 * 10668 10767: gap of 100 bp
 * 10768 11503: contig of 736 bp in length
 * 11504 11603: gap of 100 bp
 * 11604 12346: contig of 743 bp in length
 * 12347 12446: gap of 100 bp
 * 12447 13168: contig of 722 bp in length
 * 13169 13268: gap of 100 bp
 * 13269 13995: contig of 727 bp in length
 * 13996 14095: gap of 100 bp
 * 14096 14810: contig of 715 bp in length
 * 14811 14910: gap of 100 bp
 * 14911 15587: contig of 677 bp in length
 * 15588 15687: gap of 100 bp
 * 15688 16390: contig of 703 bp in length
 * 16391 16490: gap of 100 bp
 * 16491 17212: contig of 722 bp in length
 * 17213 17312: gap of 100 bp
 * 17313 18033: contig of 721 bp in length

* 18034 18033: gap of 100 bp
* 18134 18856: contig of 723 bp in length
* 18857 18956: gap of 100 bp
* 18957 19693: contig of 737 bp in length
* 19694 19793: gap of 100 bp
* 19794 20529: contig of 736 bp in length
* 20530 20629: gap of 100 bp
* 20630 21364: contig of 735 bp in length
* 21365 21464: gap of 100 bp
* 21465 22189: contig of 725 bp in length
* 22190 22289: gap of 100 bp
* 22290 22981: contig of 692 bp in length
* 22982 23081: gap of 100 bp
* 23082 23780: contig of 699 bp in length
* 23781 23880: gap of 100 bp
* 23881 24606: contig of 726 bp in length
* 24607 24706: gap of 100 bp
* 24707 25424: contig of 718 bp in length
* 25425 25524: gap of 100 bp
* 25525 26243: contig of 719 bp in length
* 26244 26343: gap of 100 bp
* 26344 27068: contig of 725 bp in length
* 27069 27168: gap of 100 bp
* 27169 27900: contig of 732 bp in length
* 27901 28000: gap of 100 bp
* 28001 28726: contig of 726 bp in length
* 28727 28926: gap of 100 bp
* 28927 29559: contig of 733 bp in length
* 29560 29659: gap of 100 bp
* 29660 30397: contig of 738 bp in length
* 30398 30497: gap of 100 bp
* 30498 31212: contig of 715 bp in length
* 31213 31312: gap of 100 bp
* 31313 32025: contig of 713 bp in length
* 32026 32125: gap of 100 bp
* 32126 32855: contig of 730 bp in length
* 32856 32955: gap of 100 bp
* 32956 33674: contig of 719 bp in length
* 33675 33774: gap of 100 bp
* 33775 34480: contig of 706 bp in length
* 34481 34580: gap of 100 bp
* 34581 35294: contig of 714 bp in length
* 35295 35394: gap of 100 bp
* 35395 36121: contig of 727 bp in length
* 36122 36221: gap of 100 bp
* 36222 36952: contig of 731 bp in length
* 36953 37052: gap of 100 bp
* 37053 37771: contig of 719 bp in length
* 37772 37871: gap of 100 bp
* 37872 38601: contig of 730 bp in length
* 38602 38701: gap of 100 bp
* 38702 39431: contig of 730 bp in length
* 39432 39531: gap of 100 bp
* 39532 40256: contig of 725 bp in length
* 40257 40356: gap of 100 bp
* 40357 41067: contig of 711 bp in length
* 41068 41167: gap of 100 bp
* 41168 41897: contig of 730 bp in length
* 41898 41997: gap of 100 bp
* 41998 42712: contig of 715 bp in length
* 42713 42812: gap of 100 bp
* 42813 43533: contig of 721 bp in length
* 43534 43633: gap of 100 bp
* 43634 44354: contig of 721 bp in length
* 44355 44454: gap of 100 bp
* 44455 45187: contig of 733 bp in length
* 45188 45287: gap of 100 bp
* 45288 46026: contig of 739 bp in length
* 46127 46126: gap of 100 bp
* 46127 46824: contig of 698 bp in length
* 46825 46924: gap of 100 bp
* 46925 47656: contig of 732 bp in length
* 47657 47756: gap of 100 bp

* 47757 48484: contig of 728 bp in length
* 48485 48584: gap of 100 bp
* 48585 49301: contig of 717 bp in length
* 49302 49401: gap of 100 bp
* 49402 50119: contig of 718 bp in length
* 50120 50219: gap of 100 bp
* 50220 50934: contig of 715 bp in length
* 50935 51034: gap of 100 bp
* 51035 51766: contig of 732 bp in length
* 51767 51866: gap of 100 bp
* 51867 52575: contig of 709 bp in length
* 52576 52675: gap of 100 bp
* 52676 53400: contig of 725 bp in length
* 53401 53500: gap of 100 bp
* 53501 54217: contig of 717 bp in length
* 54218 54317: gap of 100 bp
* 54318 55003: contig of 686 bp in length
* 55004 55103: gap of 100 bp
* 55104 55811: contig of 708 bp in length
* 55812 55911: gap of 100 bp
* 55912 56634: contig of 723 bp in length
* 56635 56734: gap of 100 bp

Query Match 100.0%; Score 231; DB 2; Length 66489;
Best Local Similarity 100.0%; Pred. No. 3.le-65;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGCGAGGCTAAGCAGTGCACAGAGATAGCTTGTGTTGGCATCTCTGTGGGAAGG 60
|||||
Db 44342 GTGGGTGCGAGGCTAAGCAGTGCACAGAGATAGCTTGTGTTGGCATCTCTGTGGGAAGG 44283

QY 61 CACGTCGTGATTGCATGAGGTTTCAGTGTCTAGTTCCTTCCCTCCAGGAAAAAGCAGACG 120
|||||
Db 44282 CACGTCGTGATTGCATGAGGTTTCAGTGTCTAGTTCCTTCCCTCCAGGAAAAAGCAGACG 44223

QY 121 GGACCTGGCCCAAGACCTCTCTGGGTGATGAGCCCTAAGGGATGTTTGTGTAGGGCCCT 180
|||||
Db 44222 GGACCTGGCCCAAGACCTCTCTGGGTGATGAGCCCTAAGGGATGTTTGTGTAGGGCCCT 44163

QY 181 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCCACCATGAGGCTTGCCT 231
|||||
Db 44162 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCCACCATGAGGCTTGCCT 44112

RESULT 13
AC105951/c

LOCUS AC105951 69090 bp DNA linear HTG 20-MAY-2002
Mus musculus clone RP23-475D21, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC105951
VERSION AC105951.2 GI:20986546
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 69090)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-475D21
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 69090)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C.,

Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

REFERENCE
AUTHORS

Submitted (11-JAN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 69090)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., FitzHugh, W., Gage, D.,
 Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L.,
 Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I.,
 Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Larocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K.,
 Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N.,
 Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J.,
 Meneses, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C.,
 Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S.,
 Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N.,
 Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,
 Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J.,
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
 Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

COMMENT

Submitted (20-MAY-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On May 20, 2002 this sequence version replaced gi:18129451.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WBIR
 Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L19149

Center clone name: 475_D_21

* NOTE: This record contains 85 individual

* sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that

* the record is updated, the accession number will

* be preserved.

* 1 714: contig of 714 bp in length

* 715 814: gap of 100 bp

* 815 1523: contig of 709 bp in length

* 1524 1623: gap of 100 bp

* 1624 2336: contig of 713 bp in length

* 2337 2436: gap of 100 bp

* 2437 3168: contig of 732 bp in length

* 3169 3268: gap of 100 bp

* 3269 3994: contig of 726 bp in length

* 3995 4094: gap of 100 bp

* 4095 4810: contig of 716 bp in length

4811 4910: gap of 100 bp

4911 5632: contig of 722 bp in length

5633 5732: gap of 100 bp

5733 6415: contig of 683 bp in length

6416 6515: gap of 100 bp

6516 7224: contig of 709 bp in length

7225 7324: gap of 100 bp

7325 8035: contig of 711 bp in length

8036 8135: gap of 100 bp

8136 8856: contig of 721 bp in length

8857 8956: gap of 100 bp

8957 9669: contig of 713 bp in length

9670 9769: gap of 100 bp

9770 10492: contig of 723 bp in length

10493 10592: gap of 100 bp

10593 11310: contig of 718 bp in length

11311 11410: gap of 100 bp

11411 12129: contig of 719 bp in length

12130 12229: gap of 100 bp

12230 12942: contig of 713 bp in length

12943 13042: gap of 100 bp

13043 13756: contig of 714 bp in length

13757 13856: gap of 100 bp

13857 14567: contig of 711 bp in length

14568 15374: contig of 707 bp in length

15375 15474: gap of 100 bp

15475 16157: contig of 683 bp in length

16158 16257: gap of 100 bp

16258 16971: contig of 714 bp in length

16972 17071: gap of 100 bp

17072 17786: contig of 715 bp in length

17787 17886: gap of 100 bp

17887 18582: contig of 696 bp in length

18583 18682: gap of 100 bp

18683 19415: contig of 733 bp in length

19416 19515: gap of 100 bp

19516 20237: contig of 722 bp in length

20238 20337: gap of 100 bp

20338 21033: contig of 696 bp in length

21034 21133: gap of 100 bp

21134 21855: contig of 722 bp in length

21856 21955: gap of 100 bp

21956 22666: contig of 711 bp in length

22667 22766: gap of 100 bp

22767 23477: contig of 711 bp in length

23478 23577: gap of 100 bp

23578 24291: contig of 714 bp in length

24292 24391: gap of 100 bp

24392 25102: contig of 711 bp in length

25103 25202: gap of 100 bp

25203 25912: contig of 710 bp in length

25913 26012: gap of 100 bp

26013 26725: contig of 713 bp in length

26726 26825: gap of 100 bp

26826 27551: contig of 726 bp in length

27552 27651: gap of 100 bp

27652 28371: contig of 720 bp in length

28372 28471: gap of 100 bp

28472 29189: contig of 718 bp in length

29190 29289: gap of 100 bp

29290 29982: contig of 693 bp in length

29983 30082: gap of 100 bp

30083 30801: contig of 719 bp in length

30802 30901: gap of 100 bp

30902 31614: contig of 713 bp in length

31615 31714: gap of 100 bp

31715 32426: contig of 712 bp in length

32427 32526: gap of 100 bp

32527 33234: contig of 708 bp in length

33235 33334: gap of 100 bp

33335 34035: contig of 701 bp in length

34036 34135: gap of 100 bp

```

* 34136 34844: contig of 709 bp in length
* 34845 34944: gap of 100 bp
* 34945 35640: contig of 696 bp in length
* 35641 35740: gap of 100 bp
* 35741 36467: contig of 727 bp in length
* 36468 36567: gap of 100 bp
* 36568 37292: contig of 725 bp in length
* 37293 37392: gap of 100 bp
* 37393 38121: contig of 729 bp in length
* 38122 38221: gap of 100 bp
* 38222 38952: contig of 731 bp in length
* 38953 39052: gap of 100 bp
* 39053 39770: contig of 718 bp in length
* 39771 39870: gap of 100 bp
* 39871 40564: contig of 694 bp in length
* 40565 40664: gap of 100 bp
* 40665 41375: contig of 711 bp in length
* 41376 41475: gap of 100 bp
* 41476 42173: contig of 698 bp in length
* 42174 42273: gap of 100 bp
* 42274 42981: contig of 708 bp in length
* 42982 43081: gap of 100 bp
* 43082 43798: contig of 717 bp in length
* 43799 43898: gap of 100 bp
* 43899 44606: contig of 708 bp in length
* 44607 44706: gap of 100 bp
* 44707 45432: contig of 726 bp in length

Query Match 100.0%; Score 231; DB 2; Length 69090;
Best Local Similarity 100.0%; Pred. No. 3e-65; 0; Indels 0; Gaps 0;
Matches 231; Conservative 0; Mismatches 0;

QY 1 GTGGGGTGCAGGCTAAGCACTGCACAGAGATAGCTTGTGTGGCATCTGTGGAAGG 60
|||||
Db 61337 GTGGGGTGCAGGCTAAGCACTGCACAGAGATAGCTTGTGTGGCATCTGTGGAAGG 61278

QY 61 CAGCTCTGATGTCATGAAGTTTCAGTCTCTAGTTCCTTCCCTCCCGAGAAACGACAG 120
|||||
Db 61277 CAGCTCTGATGTCATGAAGTTTCAGTCTCTAGTTCCTTCCCTCCCGAGAAACGACAG 61218

QY 121 GGAGCTGCCAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGGCCCT 180
|||||
Db 61217 GGAGCTGCCAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGGCCCT 61158

QY 181 ATGCTTCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 231
|||||
Db 61157 ATGCTTCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 61107

RESULT 14
AC101122/c
LOCUS Mus musculus clone RP23-93E18, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC101122
ACCESSION AC101122
VERSION AC101122.1 GI:17059896
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 70360)
Birren, B., Linton, L., Nusbaum, C. and Lander, B.
Mus musculus, clone RP23-93E18
Unpublished
REFERENCE 2 (bases 1 to 70360)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
Brown, A., Camarata, J., Campioano, A., Chang, J., Chazaro, B.,
Choepe, I., Collangelo, M., Collins, S., Collamore, A., Cook, A.,
Cooke, P., DeArelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,

```

```

Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K.,
Lamazares, R., Landers, T., Lehozky, J., Levine, R., Liu, G.,
MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L14988
Center clone name: 93_E_18
-----
* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 721: contig of 721 bp in length
* 722 821: gap of 100 bp
* 822 1528: contig of 707 bp in length
* 1529 1628: gap of 100 bp
* 1629 2332: contig of 704 bp in length
* 2333 2432: gap of 100 bp
* 2433 3144: contig of 712 bp in length
* 3145 3244: gap of 100 bp
* 3245 3967: contig of 723 bp in length
* 3968 4067: gap of 100 bp
* 4068 4791: contig of 724 bp in length
* 4792 4891: gap of 100 bp
* 4892 5599: contig of 708 bp in length
* 5600 5699: gap of 100 bp
* 5700 6425: contig of 726 bp in length
* 6426 6525: gap of 100 bp
* 6526 7219: contig of 694 bp in length
* 7220 7319: gap of 100 bp
* 7320 8037: contig of 718 bp in length
* 8038 8137: gap of 100 bp
* 8138 8846: contig of 709 bp in length
* 8847 8946: gap of 100 bp
* 8947 9664: contig of 718 bp in length
* 9665 9764: gap of 100 bp
* 9765 10491: contig of 727 bp in length
* 10492 10591: gap of 100 bp
* 10592 11316: contig of 725 bp in length
* 11317 11416: gap of 100 bp
* 11417 12126: contig of 710 bp in length
* 12127 12226: gap of 100 bp
* 12227 12943: contig of 717 bp in length
* 12944 13043: gap of 100 bp
* 13044 13755: contig of 712 bp in length

```

TITLE
JOURNAL

COMMENT

* 13756 13855: gap of 100 bp
* 13856 14559: contig of 704 bp in length
* 14560 14659: gap of 100 bp
* 14660 15404: contig of 745 bp in length
* 15405 15504: gap of 100 bp
* 15505 16228: contig of 724 bp in length
* 16229 16328: gap of 100 bp
* 16329 17033: contig of 705 bp in length
* 17034 17133: gap of 100 bp
* 17134 17838: contig of 705 bp in length
* 17839 17938: gap of 100 bp
* 17939 18664: contig of 726 bp in length
* 18665 18764: gap of 100 bp
* 18765 19475: contig of 711 bp in length
* 19476 19575: gap of 100 bp
* 19576 20305: contig of 730 bp in length
* 20306 20405: gap of 100 bp
* 20406 21141: contig of 736 bp in length
* 21142 21241: gap of 100 bp
* 21242 21965: contig of 724 bp in length
* 21966 22065: gap of 100 bp
* 22066 22804: contig of 739 bp in length
* 22805 22904: gap of 100 bp
* 22905 23637: contig of 733 bp in length
* 23638 23737: gap of 100 bp
* 23738 24472: contig of 735 bp in length
* 24473 24572: gap of 100 bp
* 24573 25300: contig of 728 bp in length
* 25301 25400: gap of 100 bp
* 25401 26122: contig of 722 bp in length
* 26123 26222: gap of 100 bp
* 26223 26935: contig of 713 bp in length
* 26936 27035: gap of 100 bp
* 27036 27770: contig of 735 bp in length
* 27771 27870: gap of 100 bp
* 27871 28550: contig of 680 bp in length
* 28551 28650: gap of 100 bp
* 28651 29358: contig of 708 bp in length
* 29359 29458: gap of 100 bp
* 29459 30183: contig of 725 bp in length
* 30184 30283: gap of 100 bp
* 30284 31012: contig of 729 bp in length
* 31013 31112: gap of 100 bp
* 31113 31829: contig of 717 bp in length
* 31830 31929: gap of 100 bp
* 31930 32642: contig of 713 bp in length
* 32643 32742: gap of 100 bp
* 32743 33463: contig of 721 bp in length
* 33464 33563: gap of 100 bp
* 33564 34280: contig of 717 bp in length
* 34281 34380: gap of 100 bp
* 34381 35089: contig of 709 bp in length
* 35090 35189: gap of 100 bp
* 35190 35922: contig of 733 bp in length
* 35923 36022: gap of 100 bp
* 36023 36739: contig of 717 bp in length
* 36740 36839: gap of 100 bp
* 36840 37552: contig of 713 bp in length
* 37553 37652: gap of 100 bp
* 37653 38375: contig of 723 bp in length
* 38376 38475: gap of 100 bp
* 38476 39195: contig of 720 bp in length
* 39196 39295: gap of 100 bp
* 39296 40019: contig of 724 bp in length
* 40020 40119: gap of 100 bp
* 40120 40844: contig of 725 bp in length
* 40845 40944: gap of 100 bp
* 40945 41660: contig of 716 bp in length
* 41661 41760: gap of 100 bp
* 41761 42445: contig of 685 bp in length
* 42446 42545: gap of 100 bp
* 42546 43241: contig of 696 bp in length
* 43242 43341: gap of 100 bp

* 43342 44062: contig of 721 bp in length
* 44063 44162: gap of 100 bp
* 44163 44878: contig of 716 bp in length
* 44879 44978: gap of 100 bp
* 44979 45691: contig of 713 bp in length
* 45692 45791: gap of 100 bp
* 45792 46522: contig of 731 bp in length
* 46523 46622: gap of 100 bp
* 46623 47356: contig of 734 bp in length
* 47357 47456: gap of 100 bp
* 47457 48159: contig of 703 bp in length
* 48160 48259: gap of 100 bp
* 48260 48989: contig of 730 bp in length
* 48990 49089: gap of 100 bp
* 49090 49803: contig of 714 bp in length
* 49804 49903: gap of 100 bp
* 49904 50616: contig of 713 bp in length
* 50617 50716: gap of 100 bp
* 50717 51436: contig of 720 bp in length
* 51437 51536: gap of 100 bp
* 51537 52267: contig of 731 bp in length
* 52268 52367: gap of 100 bp
* 52368 53073: contig of 706 bp in length
* 53074 53173: gap of 100 bp
* 53174 53887: contig of 714 bp in length
* 53888 53987: gap of 100 bp
* 53988 54714: contig of 727 bp in length
* 54715 54814: gap of 100 bp
* 54815 55551: contig of 737 bp in length
* 55552 55651: gap of 100 bp
* 55652 56375: contig of 724 bp in length
* 56376 56475: gap of 100 bp

Query Match 100.0%; Score 231; DB 2; Length 70360;

Best Local Similarity 100.0%; Pred. No. 3e-65;

Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGCGAGCTAAGCACTGCACAGAGGATAGTCTGTTCCTCCCGAGGAAAAACGACAG 60

DB 31408 GTGGGTGCGAGCTAAGCACTGCACAGAGGATAGTCTGTTCCTCCCGAGGAAAAACGACAG 31349

QY 61 CAGCTCTGATTGATGAAGTTTCAGTGTCTCCTTCCTCCCGAGGAAAAACGACAG 120

DB 31348 CAGCTCTGATTGATGAAGTTTCAGTGTCTCCTTCCTCCCGAGGAAAAACGACAG 31289

QY 121 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGGCCCT 180

DB 31288 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGGCCCT 31229

QY 181 ATGCTTGACACTGGGGATCAGACCTTACCTTCACCCATGAGGCTTGCTT 231

DB 31228 ATGCTTGACACTGGGGATCAGACCTTACCTTCACCCATGAGGCTTGCTT 31178

RESULT 15

AC123753/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

AC123753 71466 bp DNA linear HTG 01-JUN-2002
Mus musculus clone RP24-200E3, LOW-PASS SEQUENCE SAMPLING.

AC123753
AC123753.1 GI:21307315
HTG; HTGS_PHASE0.
house mouse.

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

1 (bases 1 to 71466)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.

Mus musculus, clone RP24-200E3
Unpublished

2 (bases 1 to 71466)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (01-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L26570
Center clone name: 200_E_3

* NOTE: This record contains 87 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 693 792: contig of 692 bp in length
* 693 792: gap of 100 bp
* 793 1504: contig of 712 bp in length
* 1505 1604: gap of 100 bp
* 1605 2322: contig of 718 bp in length
* 2323 2422: gap of 100 bp
* 2423 3163: contig of 741 bp in length
* 3164 3263: gap of 100 bp
* 3264 4008: contig of 745 bp in length
* 4009 4108: gap of 100 bp
* 4109 4838: contig of 730 bp in length
* 4839 4938: gap of 100 bp
* 4939 5676: contig of 738 bp in length
* 5677 5776: gap of 100 bp
* 5777 6500: contig of 724 bp in length
* 6501 6600: gap of 100 bp
* 6601 7334: contig of 734 bp in length
* 7335 7434: gap of 100 bp
* 7435 8154: contig of 720 bp in length
* 8155 8254: gap of 100 bp
* 8255 8977: contig of 723 bp in length
* 8978 9077: gap of 100 bp
* 9078 9795: contig of 718 bp in length
* 9796 9895: gap of 100 bp
* 9896 10617: contig of 722 bp in length
* 10618 10717: gap of 100 bp
* 10718 11417: contig of 700 bp in length
* 11418 11517: gap of 100 bp
* 11518 12246: contig of 729 bp in length
* 12247 12346: gap of 100 bp
* 12347 13077: contig of 731 bp in length
* 13078 13177: gap of 100 bp
* 13178 13908: contig of 731 bp in length
* 13909 14008: gap of 100 bp
* 14009 14738: contig of 730 bp in length
* 14739 14838: gap of 100 bp
* 14839 15560: contig of 722 bp in length
* 15561 15660: gap of 100 bp
* 15661 16372: contig of 712 bp in length
* 16373 16472: gap of 100 bp
* 16473 17197: contig of 725 bp in length
* 17198 17297: gap of 100 bp
* 17298 18027: contig of 730 bp in length
* 18028 18127: gap of 100 bp
* 18128 18845: contig of 718 bp in length
* 18846 18945: gap of 100 bp
* 18946 19636: contig of 711 bp in length
* 19637 19736: gap of 100 bp
* 19737 20467: contig of 711 bp in length
* 20468 20567: gap of 100 bp
* 20568 21300: contig of 733 bp in length
* 21301 21400: gap of 100 bp
* 21401 22135: contig of 735 bp in length
* 22136 22235: gap of 100 bp
* 22236 22962: contig of 727 bp in length
* 22963 23062: gap of 100 bp
* 23063 23797: contig of 735 bp in length
* 23798 23897: gap of 100 bp
* 23898 24620: contig of 723 bp in length
* 24621 24720: gap of 100 bp
* 24721 25443: contig of 723 bp in length
* 25444 25543: gap of 100 bp
* 25544 26269: contig of 726 bp in length
* 26270 26369: gap of 100 bp
* 26370 27099: contig of 730 bp in length
* 27100 27199: gap of 100 bp
* 27200 27916: contig of 717 bp in length
* 27917 28016: gap of 100 bp
* 28017 28712: contig of 696 bp in length
* 28713 28812: gap of 100 bp
* 28813 29498: contig of 686 bp in length
* 29499 29598: gap of 100 bp
* 29599 30339: contig of 741 bp in length
* 30340 30439: gap of 100 bp
* 30440 31168: contig of 729 bp in length
* 31169 31268: gap of 100 bp
* 31269 32004: contig of 736 bp in length
* 32005 32104: gap of 100 bp
* 32105 32830: contig of 726 bp in length
* 32831 32930: gap of 100 bp
* 32931 33630: contig of 700 bp in length
* 33631 33730: gap of 100 bp
* 33731 34465: contig of 735 bp in length
* 34466 34565: gap of 100 bp
* 34566 35287: contig of 722 bp in length
* 35288 35387: gap of 100 bp
* 35388 36110: contig of 723 bp in length
* 36111 36210: gap of 100 bp
* 36211 36927: contig of 717 bp in length
* 36928 37027: gap of 100 bp
* 37028 37728: contig of 701 bp in length
* 37729 37828: gap of 100 bp
* 37829 38531: contig of 693 bp in length
* 38522 38621: gap of 100 bp
* 38622 39369: contig of 748 bp in length
* 39370 39469: gap of 100 bp
* 39470 40206: contig of 737 bp in length
* 40207 40306: gap of 100 bp
* 40307 41037: contig of 731 bp in length
* 41038 41137: gap of 100 bp
* 41138 41863: contig of 726 bp in length

* 41864 41963: gap of 100 bp	
* 41964 42696: contig of 733 bp in length	
* 42697 42796: gap of 100 bp	
* 42797 43525: contig of 729 bp in length	
* 43526 43625: gap of 100 bp	
* 43626 44328: contig of 703 bp in length	
* 44329 44428: gap of 100 bp	
* 44429 45153: contig of 725 bp in length	
* 45154 45253: gap of 100 bp	
* 45254 45986: contig of 733 bp in length	
* 45987 46086: gap of 100 bp	
* 46087 46812: contig of 726 bp in length	
* 46813 46912: gap of 100 bp	
* 46913 47630: contig of 718 bp in length	
* 47631 47730: gap of 100 bp	
* 47731 48413: contig of 683 bp in length	
* 48414 48513: gap of 100 bp	
* 48514 49243: contig of 730 bp in length	
* 49244 49343: gap of 100 bp	
* 49344 50072: contig of 729 bp in length	
* 50073 50172: gap of 100 bp	
* 50173 50906: contig of 734 bp in length	
* 50907 51006: gap of 100 bp	
* 51007 51741: contig of 735 bp in length	
* 51742 51841: gap of 100 bp	
* 51842 52572: contig of 731 bp in length	
* 52573 52672: gap of 100 bp	
* 52673 53399: contig of 727 bp in length	
* 53400 53499: gap of 100 bp	
* 53500 54223: contig of 724 bp in length	
* 54224 54323: gap of 100 bp	
* 54324 55039: contig of 716 bp in length	
* 55040 55139: gap of 100 bp	
* 55140 55856: contig of 717 bp in length	
* 55857 55956: gap of 100 bp	
* 55957 56670: contig of 714 bp in length	
* 56671 56770: gap of 100 bp	
Query Match 100.0%; Score 231; DB 2; Length 71466;	
Best Local Similarity 100.0%; Pred. NO. 3e-65;	
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy 1	GTGGGGTCCGAGGCTAAGCACTGCACAGAGGATAGCTTGGCATCTGTGGGAGG 60
Db 30261	GTGGGGTCCGAGGCTAAGCACTGCACAGAGGATAGCTTGGCATCTGTGGGAGG 30202
Qy 61	CACGCTGATTCATGAAGGTTCAAGTTCCTAGTTCCTTCCCTCCAGGAAAAACGACAG 120
Db 30201	CACGCTGATTCATGAAGGTTCAAGTTCCTAGTTCCTTCCCTCCAGGAAAAACGACAG 30142
Qy 121	GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCCTAAGGGATGTTTGTAGGGCCCT 180
Db 30141	GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCCTAAGGGATGTTTGTAGGGCCCT 30082
Qy 181	ATGCTTGCACACTGGGATCAGACCTTACCTTACCCATGAGGCTTGCCT 231
Db 30081	ATGCTTGCACACTGGGATCAGACCTTACCTTACCCATGAGGCTTGCCT 30031

Search completed: December 24, 2002, 13:23:27
Job time : 2788 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen, Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 12:22:46 ; Search time 255 Seconds
(without alignments)
2040.046 Million cell updates/sec

Title: US-09-673-716-1

Perfect score: 231

Sequence: 1 ggggggtaggaggaagca.....ttcaccatgaggtgctt 231

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Genesec_101002.*

1: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/genesec/genesecq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/genesec/genesecq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	231	100.0	231	21	AAZ55983
2	231	100.0	393	21	AAZ55986
3	224	97.0	226	21	AAZ55987
4	89.4	38.7	5018	6	AA50150
5	89.4	38.7	5018	8	AA70646
6	87.8	38.0	3336	6	AA50149
7	86.2	37.3	1434	18	AAV04900
8	81.4	35.2	1400	18	AAV04901
9	35	15.2	50	16	AAQ85497
Core fragment B, a murine genomic core PRE region.					
Post translational regulatory element; PRE; murine; NCTE;					
Post-transcriptional RNA nucleo-cytoplasmic transport element;					
attenuated virus; retrovirus; HIV-1; Intracisternal A-type particle; IAP;					
vaccine; viral infection; core fragment B; ds.					
Mus sp.					
WO9961596-A2.					
02-DEC-1999.					
18-MAY-1999; 99WO-US11082.					
22-MAY-1998; 98US-0086487.					
(USSH) US DEPT HEALTH & HUMAN SERVICES.					
Pavlakis GN, Nappi F;					
WPI; 2000-116336/10.					
Novel post-transcriptional regulatory elements used to construct					

Human histone deac
Human ovarian can
Human colon cancer
Murine post transi
DNA encoding a hum
Human DNA repair a
EST with homology
Expressed sequence
Human full-length
Human fibrillin 3
Human fibrillin 3
Human cDNA differe
Arabidopsis thalia
Arabidopsis thalia
DNA encoding human
DNA encoding human
Human colon cancer
Genomic sequence #
P. chrysogenum sut
Genomic P. chrysog
Sequence encoding
Drosophila melanog
Drosophila melanog
Human DNA sequence
Human DNA sequence
Novel human coding
N. meningitidis pa
Neisseria meningit
Murine post transi
Human gene express
Human immune/haema
Genomic sequence #
Drosophila melanog
Drosophila melanog
DNA encoding novel
DNA encoding novel

ALIGNMENTS

RESULT 1
AAZ55983
ID AAZ55983 standard; DNA; 231 BP.

AC AAZ55983;

DT 10-APR-2000 (first entry)

DE Core fragment B, a murine genomic core PRE region.

KW Post translational regulatory element; PRE; murine; NCTE;

KW post-transcriptional RNA nucleo-cytoplasmic transport element;

KW attenuated virus; retrovirus; HIV-1; Intracisternal A-type particle; IAP;

KW vaccine; viral infection; core fragment B; ds.

OS Mus sp.

PN WO9961596-A2.

XX 02-DEC-1999.

XX 18-MAY-1999; 99WO-US11082.

XX 22-MAY-1998; 98US-0086487.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Pavlakis GN, Nappi F;

XX WPI; 2000-116336/10.

XX Novel post-transcriptional regulatory elements used to construct

in Vlnke

PT attenuated retroviruses for vaccines -
 FS Claim 3; Page 44; 59pp; English.
 CC This sequence represents a core post-transcriptional regulatory
 CC element (PRE) region, designated core fragment B, which is located
 CC within PRE7 (AA255986). Core fragment B can function as a post-
 CC transcriptional RNA nucleocytoplasmic transport element (NCTE). The PRE
 CC is used to replace the NCTE of HIV-1, producing a virus with lower
 CC replicative activity, resulting in an attenuated virus. This novel PRE
 CC was initially derived from a murine genomic sequence and sequence
 CC analysis found that it had significant homology to intracisternal A-type
 CC particle (IAP) sequences. The nucleic acids and oligonucleotides of the
 CC invention can be delivered into cells cultures, tissues and organisms
 CC for synthesis, mutation and screening. When the PRE is used to replace
 CC the NCTE of viruses, especially retroviruses such as HIV, an attenuated
 CC virus is produced, which may be used in a viral vaccine for the
 CC prophylaxis or amelioration of a viral infection in a mammal.
 XX Sequence 231 BP; 47 A; 58 C; 70 G; 56 T; 0 other;
 SQ Query Match 100.0%; Score 231; DB 21; Length 231;
 Best Local Similarity 100.0%; Pred. No. 2.7e-70;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGGGGTCGAGGCTAGCAGATGCACAGAGATAGCTTGTGGCATCTGTGGGAGG 60
 DB 1 GTGGGGTCGAGGCTAGCAGATGCACAGAGATAGCTTGTGGCATCTGTGGGAGG 60
 QY 61 CACCTCTGATTGCATGAGGTTCAAGTTCCTAGTTCCTTCCCTCCAGGAAACGACACG 120
 DB 61 CACGCTCTGATTGCATGAGGTTCAAGTTCCTAGTTCCTTCCCTCCAGGAAACGACACG 120
 QY 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTAAGGGATGGTTTGTGAGGGCCCT 180
 DB 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTAAGGGATGGTTTGTGAGGGCCCT 180
 QY 181 ATGCTTGACACTGGGATGATGAGCCTTACCTTCCACCATGAGGCTTGCCTT 231
 DB 181 ATGCTTGACACTGGGATGATGAGCCTTACCTTCCACCATGAGGCTTGCCTT 231
 RESULT 2
 AA255986
 ID AA255986 standard; DNA; 393 BP.
 AC AA255986;
 DT 10-APR-2000 (first entry)
 XX Murine genomic post-transcriptional regulatory element, PRE7.
 DE Post translational regulatory element; PRE; murine; NCTE;
 KW post-transcriptional RNA nucleocytoplasmic transport element;
 KW attenuated virus; retrovirus; HIV-1; intracisternal A-type particle; IAP;
 KW vaccine; viral infection; PRE7; ds.
 XX Mus sp.
 OS WO9961596-A2.
 PN 02-DEC-1999.
 PD 18-MAY-1999; 99WO-US11082.
 XX 22-MAY-1998; 98US-0086487.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Pavlakis GN, Nappi F;
 PI WPI; 2000-116336/10.
 XX

inv

PT Novel post-transcriptional regulatory elements used to construct
 PT attenuated retroviruses for vaccines -
 XX Example 1; Page 44; 59pp; English.
 CC The invention relates to the core region of a novel post-transcriptional
 CC regulatory element (PRE), designated core fragment B (AA255983), that
 CC can function as a post-transcriptional RNA nucleocytoplasmic transport
 CC element (NCTE). The PRE is used to replace the NCTE of HIV-1, producing
 CC a virus with lower replicative activity, resulting in an attenuated
 CC virus. This novel PRE was initially derived from a murine genomic
 CC sequence and sequence analysis found that it had significant homology to
 CC intracisternal A-type particle (IAP) sequences. The nucleic acids and
 CC oligonucleotides of the invention can be delivered into cells cultures,
 CC tissues and organisms for synthesis, mutation and screening. When the
 CC PRE is used to replace the NCTE of viruses, especially retroviruses such
 CC as HIV, an attenuated virus is produced, which may be used in a viral
 CC vaccine for the prophylaxis or amelioration of a viral infection in a
 CC mammal. This sequence represents PRE7, a murine sequence which comprises
 CC core fragment B which was characterised in an exemplification of the
 CC present invention, and is functional as a PRE.
 XX Sequence 393 BP; 85 A; 97 C; 109 G; 102 T; 0 other;
 SQ Query Match 100.0%; Score 231; DB 21; Length 393;
 Best Local Similarity 100.0%; Pred. No. 3.3e-70;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGGGGTCGAGGCTAGCAGATGCACAGAGATAGCTTGTGGCATCTGTGGGAGG 60
 DB 99 GTGGGGTCGAGGCTAGCAGATGCACAGAGATAGCTTGTGGCATCTGTGGGAGG 158
 QY 61 CACCTCTGATTGCATGAGGTTCAAGTTCCTAGTTCCTTCCCTCCAGGAAACGACACG 120
 DB 159 CACGCTCTGATTGCATGAGGTTCAAGTTCCTAGTTCCTTCCCTCCAGGAAACGACACG 218
 QY 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTAAGGGATGGTTTGTGAGGGCCCT 180
 DB 219 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTAAGGGATGGTTTGTGAGGGCCCT 278
 QY 181 ATGCTTGACACTGGGATGATGAGCCTTACCTTCCACCATGAGGCTTGCCTT 231
 DB 279 ATGCTTGACACTGGGATGATGAGCCTTACCTTCCACCATGAGGCTTGCCTT 329
 RESULT 3
 AA255987
 ID AA255987 standard; DNA; 226 BP.
 XX AA255987;
 AC AA255987;
 DT 10-APR-2000 (first entry)
 XX Post-transcriptional regulatory element PRE7 functional fragment, M4.
 DE Post translational regulatory element; PRE; murine; NCTE;
 KW post-transcriptional RNA nucleocytoplasmic transport element;
 KW attenuated virus; retrovirus; HIV-1; intracisternal A-type particle; IAP;
 KW vaccine; viral infection; M4; ds.
 XX Unidentified.
 OS WO9961596-A2.
 PN 02-DEC-1999.
 PD 18-MAY-1999; 99WO-US11082.
 XX 22-MAY-1998; 98US-0086487.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 PA Pavlakis GN, Nappi F;
 PI

XX WPI; 2000-116336/10.

XX Novel post-transcriptional regulatory elements used to construct

PT attenuated retroviruses for vaccines

XX

XX Example 1; Page 45; 59pp; English.

XX The invention relates to the core region of a novel post-transcriptional

CC regulatory element (PRE), designated core fragment B (AAZ55983), that

CC can function as a post-transcriptional RNA nucleocytoplasmic transport

CC element (NCTE). The PRE is used to replace the NCTE of HIV-1, producing

CC a virus with lower replicative activity, resulting in an attenuated

CC virus. This novel PRE was initially derived from a murine genomic

CC sequence and sequence analysis found that it had significant homology to

CC intracellular A-type particle (IAP) sequences. The nucleic acids and

CC oligonucleotides of the invention can be delivered into cells cultures,

CC tissues and organisms for synthesis, mutation and screening. When the

CC PRE is used to replace the NCTE of viruses, especially retroviruses such

CC as HIV, an attenuated virus is produced, which may be used in a viral

CC vaccine for the prophylaxis or amelioration of a viral infection in a

CC mammal. This sequence represents M4, a sequence which is very similar to

CC core fragment B (having 7 additional nucleotides at the 3' end and 2

CC fewer nucleotides at the 5' end). M4 was characterised in an

CC exemplification of the present invention, and is functional as a PRE.

XX

XX Sequence 226 BP; 47 A; 58 C; 69 G; 52 T; 0 other;

SQ

Query Match 97.0%; Score 224; DB 21; Length 226;

Best Local Similarity 100.0%; Pred. No. 7.4e-68;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGCGAGGCTAAGCACTGCACAGAGGATAGCTGTGCTGGCATCTGTGGAAGG 60

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 3 GTGGGTGCGAGGCTAAGCACTGCACAGAGGATAGCTGTGCTGGCATCTGTGGAAGG 62

QY 61 CAGCTCTGATTGTCAGAGGTTTCAGTGTCTCTAGTTCCTTCCCTCCAGGAAACGACACG 120

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 63 CAGCTCTGATTGTCAGAGGTTTCAGTGTCTCTAGTTCCTTCCCTCCAGGAAACGACACG 122

QY 121 GGAGCTGGCCAGACCTCTCTGGGTGATGAGCCCTAAGGATGTTTGTGTAGGGCCCTT 180

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 123 GGAGCTGGCCAGACCTCTCTGGGTGATGAGCCCTAAGGATGTTTGTGTAGGGCCCTT 182

QY 181 ATGCTTGACACATGGGGATCAGACCTCTACCTTCACCCATGAGG 224

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 183 ATGCTTGACACATGGGGATCAGACCTCTACCTTCACCCATGAGG 226

RESULT 4

AAAN50150

ID AAAN50150 standard; cDNA; 5018 BP.

XX

XX AAAN50150;

XX

XX 16-AUG-2002 (updated)

DT 27-SEP-1991 (first entry)

XX

XX Sequence of the cDNA clone 23B6p10.2 encoding a polypeptide

DE exhibiting mammalian immunoglobulin binding factor activity (IBF).

DE

XX Immunoglobulin E-mediated disease; therapy; B-cell differentiation;

KW Immunoglobulin E-binding factor; ss.

KW

XX Rattus sp.

OS

XX

XX Key Location/Qualifiers

FT CDS 423..2750

FT /*tag= a

XX

XX EP155192-A.

PN

XX 18-SEP-1985.

PD

XX 15-MAR-1985; 85EP-0301834.

XX

XX 16-MAR-1984; 84US-0590430.

XX

XX (SCHE-) SCHERING BIOTECH CO.

PA (UYJO) JOHNS HOPKINS UNIV.

PA (SCHE) SCHERING-BIOTECH CO.

PA (DNAX-) DNAX RES INST MOLEC.

XX

XX Martens CL, Ishizaka K, Moore KW, Huff TF;

PI WPI; 1985-231863/38.

DR P-PSDB; AAP50121.

DR

XX New complementary DNA clones coding for poly:peptide(s) - with

PT sequence of mammalian immunoglobulin factor and obtd. from

PT transformed or transfected host

XX

XX Claim 8; Page 59-64; 71pp; English.

XX

XX IBF and IBF cDNA are useful in studies on the immune system.

CC Treatment for IgE-mediated diseases may be possible. IBF may enhance

CC B-cell differentiation into an immunoglobulin-secreting cell.

CC (Updated on 16-AUG-2002 to add missing OS field.)

XX

XX Sequence 5018 BP; 1424 A; 1086 C; 1165 G; 1343 T; 0 other;

SQ

Query Match 38.7%; Score 89.4; DB 6; Length 5018;

Best Local Similarity 74.9%; Pred. No. 1.4e-20;

Matches 128; Conservative 0; Mismatches 36; Indels 7; Gaps 1;

QY 55 GGAAGGCACCTCGATTGTCATGAGGTTTCAGTGTCTCTAGTTCCTTCCCTCCAGGCAAAAC 114

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 4449 GGGAGACATGTCATCTTTCATGAGGTTTCAGTGTCTCTAGTTCCTTCCCTCCAGGCAAAAC 4508

QY 115 GACACGGGAGCTGGCCAGACCTCTCTGGGTGA-----TGACCCTAAGGATGGTTT 167

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 4509 GACACGGGAGGAGCTGAGGTTGCTCTGGGTAAAGCCTGTAGACCTTAAGAGCTTAATCCT 4568

QY 168 GTGTAGGGCCCCCTATGCTTGCACACTGGGATCAGACCTTACCTTCACCC 218

DB | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

DB 4569 GTACATGGCTCCCTTACCTACACACTGGGATTTGACCTCTATCTCCACTC 4619

RESULT 5

AAAN70646

ID AAAN70646 standard; cDNA; 5018 BP.

XX

XX AAAN70646;

XX

XX 27-APR-1991 (first entry)

DT

XX Clone 23B6p102 encoding polypeptide with IgE binding factor

DE activity.

DE

XX Antibody; Immunoglobulin; ss.

KW

XX Key Location/Qualifiers

FT CDS 423..2753

FT /*tag= a

XX

XX JP62045600-A.

PN

XX

XX 27-FEB-1987.

PD

XX

XX 21-AUG-1985; 85JP-0183810.

PF

XX 21-AUG-1985; 85JP-0183810.

PR

XX (SCHE) SCHERING BIOTECH CO.

PA

XX WPI; 1987-09765/14.

DR

DR P-PSDB; AAP70417.
 XX CDNA clone coding polypeptide - showing IgE bond factor activity.
 XX PS Disclosure; Page 29; 33pp; Japanese.
 XX The clone encodes a polypeptide which shows IgE binding factor
 CC activity. See also AAN70645.
 XX Sequence 5018 BP; 1423 A; 1087 C; 1165 G; 1343 T; 0 other;
 SQ

Query Match 38.7%; Score 89.4; DB 8; Length 5018;
 Best Local Similarity 74.9%; Pred. No. 1.4e-20;
 Matches 128; Conservative 0; Mismatches 36; Indels 7; Gaps 1;
 QY 55 GGAAGGCACTGCTGATTCATGAAGGTTCAAGTTCAGTTCCTCCCTCCCGAGGAAAC 114
 DB 4449 GGGAGACATGATCTTCATGAAGGTTCAAGTTCAGTTCCTCCCTCCCGAGGAAAC 4508
 QY 115 GACAGGAGCTGGCCAGACCTCTCTGGGTGA-----TGAGCCTAAGGGATGGTTT 167
 DB 4509 GACAGGAGCAGGTTCAGGTTGCTCTGGGTAAAGCCTGTAAGCCTAAGAGCTAATCCT 4568
 QY 168 GTGAGGCCCCCTATGCTGTCACACTGGGGATCAGACCTCTACCTTCACCC 218
 DB 4569 GTACATGGCTCTTACCTACACACTGGGGATTTGAOCCTCTATCTCCACTC 4619

RESULT 6
 ID AAN50149 standard; cDNA; 3336 BP.
 AC AAN50149;
 XX
 DT 16-AUG-2002 (updated)
 DT 27-SEP-1991 (first entry)
 XX
 DE Sequence of the cDNA clone 23B6p8.3 encoding a polypeptide
 DE exhibiting mammalian immunoglobulin binding factor activity (IBF).
 XX
 KW Immunoglobulin E-mediated disease; therapy; B-cell differentiation;
 KW immunoglobulin E-binding factor; ss.
 XX
 OS Rattus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 94..1767
 FT /*tag= a
 XX
 FN EP155192-A.
 XX
 PD 18-SEP-1985.
 XX
 PF 15-MAR-1985; 85EP-0301834.
 XX
 PR 16-MAR-1984; 84US-0590430.
 XX
 PA (SCHE-) SCHERING BIOTECH CO.
 PA (UJO) JOHNS HOPKINS UNIV.
 PA (SCHE) SCHERING-BIOTECH CO.
 PA (DNAX-) DNAX RES INST MOLEC.
 XX
 PI Martens CL, Ishizaka K, Moore KW, Huff TF;
 XX
 XX WPI; 1985-231863/38.
 DR P-PSDB; AAP50120.
 XX
 XX New complementary DNA clones coding for poly:peptide(s) - with
 PT sequence of mammalian immunoglobulin factor and obtd. from
 PT transformed or transfected host
 XX
 XX Claim 8; Page 55-58; 71pp; English.
 PS
 XX

CC IBF and IBF cDNA are useful in studies on the immune system.
 CC Treatment for IGE-mediated diseases may be possible. IBF may enhance
 CC B-cell differentiation into an immunoglobulin-secreting cell.
 CC (Updated on 16-AUG-2002 to add missing OS field.)
 XX
 SQ Sequence 3336 BP; 944 A; 725 C; 839 G; 828 T; 0 other;
 Query Match 38.0%; Score 87.8; DB 6; Length 3336;
 Best Local Similarity 74.3%; Pred. No. 4.3e-20;
 Matches 127; Conservative 0; Mismatches 37; Indels 7; Gaps 1;
 QY 55 GGAAGGACGCTGATTCATGAAGGTTCAAGTTCAGTTCCTCCCTCCCGAGGAAAC 114
 DB 2772 GGAACATGTCATCTTTTCATGAAGGTTCAAGTTCAGTTCCTCCCTCCCGAGGAAAC 2831
 QY 115 GACAGGAGCTGCCAAGACCTCTCTGGGTGA-----TGAGCCTAAGGGATGGTTT 167
 DB 2832 GACAGGAGCAGGTTCAGGTTGCTCTGGGTAAAGCCTGTAAGCCTAAGAGCTAATCCT 2891
 QY 168 GTGAGGCCCCCTATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCC 218
 DB 2892 GTACATGGCTCTTACCTACACACTGGGGATTTGACCTCTATCTCCACTC 2942

RESULT 7
 ID AAV04900 standard; DNA; 1434 BP.
 AC AAV04900;
 XX
 DT 30-APR-1998 (first entry)
 XX
 DE Nucleotide sequence of a portion of a neo-minichromosome.
 XX
 KW Neo-minichromosome; multicentric; dicentric; chromosome; euchromatin;
 KW heterochromatic DNA; minichromosome; artificial chromosome;
 KW selectable marker; satellite artificial chromosome; SATAC; genomic locus;
 KW targeted integration; transgenic animal; therapeutic product;
 XX
 KW gene therapy; cloning vehicle; genomic DNA library; ss.
 XX
 OS Mus sp.
 XX
 PN WO9740183-A2.
 XX
 PD 30-OCT-1997.
 XX
 PF 10-APR-1997; 97WO-US05911.
 XX
 PR 07-AUG-1996; 96US-0695191.
 PR 10-APR-1996; 96US-0629822.
 PR 15-JUL-1996; 96US-0682080.
 XX
 PA (AMGE-) AMERICAN GENE THERAPY INC.
 PA (BIOL-) BIOLOGICAL RES CENT HUNGARIAN ACAD SCI.
 PA (UYLO-) UNIV LOMA LINDA.
 XX
 PI Hadlaczky G, Szalay AA;
 XX
 DR WPI; 1997-535860/49.
 XX
 PT Producing satellite artificial chromosomes or mini:chromosomes -
 PT useful for, e.g. cloning multiple proteins of a metabolic pathway or
 PT multivalent vaccines, etc.
 XX
 PS Claim 85; Page 203; 248pp; English.
 XX
 CC Nucleotide sequences AAV04900-902 are portions of a neo-minichromosome
 CC isolated from a mouse rRNA gene. 3 products were obtained from a single
 CC amplification reaction, which suggests that the sequence of the DNA
 CC located between different sets of inverted repeats may differ.
 CC AAV04900-01 show high (96%) sequence homology to portions of DNA from
 CC intracisternal A-particles from mouse. The minichromosome is derived from
 CC multicentric, typically dicentric, chromosome that contains more

CC euchromatin than heterochromatic DNA. The minichromosome is an example
 CC of an artificial chromosome. Artificial chromosomes can be produced by
 CC incorporating a DNA fragment comprising a selectable marker into a cell's
 CC genomic DNA, into or adjacent to an amplifiable region, and selecting a
 CC cell that comprises either a satellite artificial chromosome (SATAC)
 CC (see AAV04903-09), a de novo centromere, or a minichromosome (e.g.
 CC present sequence). Artificial chromosomes provide an extra genomic locus
 CC for targeted integration of megabase size DNA fragments that contain
 CC single or multiple genes. SATACs can be introduced into embryonic cells
 CC of non-human animals to produce transgenic animals that express a
 CC heterologous DNA encoding a therapeutic product, e.g. anti-HIV. The DNA
 CC may also encode antigens that induce immunoprotective response against
 CC pathogen. These therapeutic products can be used in gene therapy. The
 CC artificial chromosomes are useful as cloning vehicles that accommodate
 CC entire genomes for the preparation of genomic DNA libraries, and also
 CC for the production of proteins which may be involved in a biochemical
 CC pathway or in multivalent vaccines.
 XX
 SQ Sequence 1434 BP; 352 A; 320 C; 344 G; 418 T; 0 other;
 Query Match 37.3%; Score 86.2; DB 18; Length 1434;
 Best Local Similarity 73.7%; Pred. No. 1.2e-19;
 Matches 126; Conservative 0; Mismatches 38; Indels 7; Gaps 1;
 QY 55 GGAAGGCAGCTCTGATTGCATGAAGCTTCAGTGTCTTCCCTTCCCTCCCGAGGAAAAAC 114
 Db 900 GGGGGACATGTCATCTTTTCATGAAGGTTTCAGTGTCTTCCCTTCCCTCCCGAGGAAAAAC 959
 QY 115 GACACGGGAGCTGGCCCAAGACCTCTCTGGGTGA-----TGAGCCTTAAGGGATGGTTT 167
 Db 960 GACACGGGAGCAGGTGAGGTTGCTCTGGGTAAAGCCTGTGAGCCTGGGAGCTATTCCT 1019
 QY 168 GTGTAGGGCCCTTATGCTTGCACACTGGGGATCAGACCTCTTACCTTCACCC 218
 Db 1020 GTACATGGCTCTTTTACCTACACACTGGGGATTGACCTCTATCTCCACTC 1070
 RESULT 8
 AAV04901
 ID AAV04901 standard; DNA; 1400 BP.
 XX
 AC AAV04901;
 XX
 DT 30-APR-1998 (first entry)
 XX
 DE Nucleotide sequence of a portion of a neo-minichromosome.
 XX Neo-minichromosome; multicentric; dicentric; chromosome; euchromatin;
 KW heterochromatic DNA; minichromosome; artificial chromosome;
 KW selectable marker; satellite artificial chromosome; SATAC; genomic locus;
 KW targeted integration; transgenic animal; therapeutic product;
 KW gene therapy; cloning vehicle; genomic DNA library; ss.
 XX
 OS Mus sp.
 XX
 PN WO9740183-A2.
 XX
 PD 30-OCT-1997.
 XX
 PF 10-APR-1997; 97WO-US05911.
 XX
 PR 07-AUG-1996; 96US-0695191.
 PR 10-APR-1996; 96US-0629822.
 PR 15-JUL-1996; 96US-0682080.
 XX
 PA (AMGE-) AMERICAN GENE THERAPY INC.
 PA (BIOL-) BIOLOGICAL RES CENT HUNGARIAN ACAD SCI.
 PA (UYLO-) UNIV LOMA LINDA.
 XX
 PI Hladacsky G, Szalay AA;
 XX
 DR WPI; 1997-535860/49.
 XX

PT Producing satellite artificial chromosomes or mini:chromosomes -
 PT useful for, e.g. cloning multiple proteins of a metabolic pathway or
 PT multivalent vaccines, etc.
 XX
 PS Claim 85; Page 204; 248pp; English.
 XX
 CC Nucleotide sequences AAV04900-902 are portions of a neo-minichromosome
 CC isolated from a mouse rRNA gene. 3 products were obtained from a single
 CC amplification reaction, which suggests that the sequence of the DNA
 CC located between different sets of inverted repeats may differ.
 CC AAV04900-01 show high (96%) sequence homology to portions of DNA from
 CC intracisternal A-particles from mouse. The minichromosome is derived from
 CC multicentric, typically dicentric, chromosome that contains more
 CC euchromatin than heterochromatic DNA. The minichromosome is an example
 CC of an artificial chromosome. Artificial chromosomes can be produced by
 CC incorporating a DNA fragment comprising a selectable marker into a cell's
 CC genomic DNA, into or adjacent to an amplifiable region, and selecting a
 CC cell that comprises either a satellite artificial chromosome (SATAC)
 CC (see AAV04903-09), a de novo centromere, or a minichromosome (e.g.
 CC present sequence). Artificial chromosomes provide an extra genomic locus
 CC for targeted integration of megabase size DNA fragments that contain
 CC single or multiple genes. SATACs can be introduced into embryonic cells
 CC of non-human animals to produce transgenic animals that express a
 CC heterologous DNA encoding a therapeutic product, e.g. anti-HIV. The DNA
 CC may also encode antigens that induce immunoprotective response against
 CC pathogen. These therapeutic products can be used in gene therapy. The
 CC artificial chromosomes are useful as cloning vehicles that accommodate
 CC entire genomes for the preparation of genomic DNA libraries, and also
 CC for the production of proteins which may be involved in a biochemical
 CC pathway or in multivalent vaccines.
 XX
 SQ Sequence 1400 BP; 341 A; 310 C; 345 G; 404 T; 0 other;
 Query Match 35.2%; Score 81.4; DB 18; Length 1400;
 Best Local Similarity 71.9%; Pred. No. 5.4e-18;
 Matches 123; Conservative 0; Mismatches 41; Indels 7; Gaps 1;
 QY 55 GGAAGGCAGCTCTGATTGCATGAAGTTCAGTGTCTTCCCTTCCCGAGGAAAAAC 114
 Db 771 GGGAGACATGTCATCTTTTCAAGAAGGTTGAGTGTCCAAGTGTCTTCCCTCCAGGCAAAAC 830
 QY 115 GACACGGGAGCTGGCCCAAGACCTCTCTGGGTGA-----TGAGCCTTAAGGGATGGTTT 167
 Db 831 GACACGGGAGCAGGTGAGGTTGCTCTGGGTAAAGCCTGTGAGCCTTAAGAGCTATTCCT 890
 QY 168 GTGTAGGGCCCTTATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCC 218
 Db 891 GTACATGGCTCTTTCCTTACCTACACACTGGGGATTGACCTCTATCTCCACTC 941
 RESULT 9
 AAV04901
 ID AAV04901 standard; DNA; 50 BP.
 XX
 AC AAV04901;
 XX
 DT 10-JAN-1996 (first entry)
 XX
 DE DNA probe 3 detects DNA-protein complex in immortal cells.
 XX
 KW DNA-protein complex; detection; proliferation; tumour formation;
 KW diagnose; malignancy; biopsy; probe; ss.
 XX
 OS Synthetic.
 XX
 PN WO9502701-A1.
 XX
 PD 26-JAN-1995.
 XX
 PF 13-JUL-1994; 94WO-EP02307.
 XX
 PR 15-JUL-1993; 93DE-4323727.
 XX

(BOEUF) BOEHRINGER MANNHEIM GMBH.

Abken H, Albert W, Jungfer H, Abken HJ;
WPI; 1995-067344/09.

New DNA-protein complex characteristic of cells with unlimited proliferation capacity - and its components and derived antibodies, useful in diagnosing malignant tumours.

Claim 1; Page 51; 106pp; German.

The DNA shown is found in human or animal cells that have an unlimited capacity for unlimited cell proliferation or tumour formation. They have no ability to promote immortalisation of the cells, and are usually found in a DNA-protein complex in the cell cytoplasm. The DNA is useful for detection of these complexes, and diagnosis of malignant tumours.

Differentiation between cells with unlimited and only transient cell proliferation is possible. (See also AAQ85493-54)

Sequence 50 BP; 13 A; 15 C; 9 G; 13 T; 0 other;

Query Match 15.2%; Score 35; DB 16; Length 50;
Best Local Similarity 100.0%; Pred. NO. 0.022;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

197 GATCAGACCTCTACTTCACCATGAGGCTTGCTT 231
|||||
1 GATCAGACCTCTACTTCACCATGAGGCTTGCTT 35

RESULT 10
AAC89560/C
ID AAC89560 standard; DNA; 122186 BP.
XX
AC AAC89560;
XX
08-MAR-2001 (first entry)
XX
Human histone deacetylase HDAC-D coding sequence.
DE
XX
XX
KW Histone deacetylase; HDAC-1; HDAC-2; HDAC-3; HDAC-4; HDAC-5; HDAC-C;
KW HDAC-D; cell cycle; tumorigenesis; cancer; inhibitor; antisense;
KW gene therapy; ds.
XX
OS Homo sapiens.
XX
XX WO200071703-A2.
XX
XX 30-NOV-2000.
XX
XX 03-MAY-2000; 2000WO-IB01252.
XX
XX 03-MAY-1999; 99US-0132287.
XX
XX (METH-) METHYLGENE INC.
XX
XX Macleod AR, Li Z, Besterman JM;
PI
XX
XX WPI; 2001-016407/02.
XX
XX
XX Antisense oligonucleotide that inhibits expression of a histone deacetylase, useful for treating and/or alleviating the symptoms of neoplasia, or for inhibiting neoplastic cell growth in an animal -
PT
PT
PT
XX
XX Disclosure; Page 89-125; 125pp; English.

The present invention provides inhibitors of histone deacetylase enzymes such as HDAC-1, HDAC-2, HDAC-3, HDAC-4, HDAC-5, HDAC-C and HDAC-D. These inhibitors may be antisense strands or they may be compounds identified by contacting the enzyme with the compound and measuring the identified enzyme activity. These inhibitors are useful for treating cancers and for identifying which histone deacetylase is involved in a neoplasia.

of the differentially expressed gene product is correlated with a cancerous state of the cell from which the test sample was derived. The polynucleotide sequences can be used in a method for detecting differentially expressed genes correlated with a cancerous state of a mammalian cell. The polynucleotides can also be used as probes for detecting and mapping related genes. They can be used in diagnosis and prognosis of diseases and disorders (e.g. identification of pre-metastatic or metastatic cancerous states, stages of cancer, or

```

Query Match      13.0%; Score 30; DB 21; Length 30;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGGTCCGAGGCTAAGCACTGCACAGAG 30
    |||||
Db 1 GTGGGGTCCGAGGCTAAGCACTGCACAGAG 30

RESULT 14
AAX61423/c
ID AAX61423 standard; cDNA; 610 BP.
XX
AC AAX61423;
XX
DT 14-JUL-1999 (first entry)
XX
DE DNA encoding a human secreted protein.
XX
KW Human secreted protein; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; fetal deficiency; blood disorder; leukemia;
KW immune system disease; autoimmune disease; hepatic disease;
KW renal disease; lymphoma; inflammation; allergy; ischemic shock;
KW Alzheimer's; cognitive disorder; schizophrenia; prostate disease;
KW obesity; osteoclast; osteoporosis; arthritis; malignancy; testes disease;
KW lung disease; thymus disease; digestive disorder; endocrine disorder;
KW infection; AIDS; ss.
XX
OS Homo sapiens.
XX
PN WO9922243-A1.
XX
PD 06-MAY-1999.
XX
XX
XX 23-OCT-1998; 98WO-US22376.
XX
XX 24-OCT-1997; 97US-0063387.
XX
XX 24-OCT-1997; 97US-0062784.
XX
XX 24-OCT-1997; 97US-0063088.
XX
XX 24-OCT-1997; 97US-0063089.
XX
XX 24-OCT-1997; 97US-0063090.
XX
XX 24-OCT-1997; 97US-0063091.
XX
XX 24-OCT-1997; 97US-0063092.
XX
XX 24-OCT-1997; 97US-0063097.
XX
XX 24-OCT-1997; 97US-0063098.
XX
XX 24-OCT-1997; 97US-0063099.
XX
XX 24-OCT-1997; 97US-0063100.
XX
XX 24-OCT-1997; 97US-0063101.
XX
XX 24-OCT-1997; 97US-0063109.
XX
XX 24-OCT-1997; 97US-0063110.
XX
XX 24-OCT-1997; 97US-0063111.
XX
XX 24-OCT-1997; 97US-0063148.
XX
XX 24-OCT-1997; 97US-0063386.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Brewer LA, Carter KC, Duan DR, Ebner R, Endress GA;
XX Feng P, Florence C, Florence KA, Greene JM, Janat F;
XX Kay H, Lafleur DW, Moore PA, Ni J, Olsen HS, Rosen CA;
XX Ruben SM, Shi Y, Soppet DR, Wei Y, Young P;
XX
XX WPI; 1999-303069/25.
XX
XX P-PSDB; AAY19543.
XX
XX
XX New isolated human genes and the secreted polypeptides they encode
XX
XX Claim 3; Page 363; 546pp; English.
XX
XX The specification describes cDNA sequences (AAX61322-X61470) encoding
XX human secreted proteins (AAY19442-Y19590). The polynucleotides and their
XX corresponding secreted polypeptides are useful for preventing, treating
XX or ameliorating medical conditions, e.g. by protein or gene therapy.
XX Pathological conditions can also be diagnosed by determining the amount

```

```

of the polypeptides in a sample or by determining the presence of
mutations in the polynucleotides. Specific uses are described for each
of the polynucleotides, based on which tissues they are most highly
expressed in, and include developing products for the diagnosis or
treatment of cancer, tumours, neurodegenerative disorders, developmental
abnormalities and fetal deficiencies, blood disorders, leukemias,
diseases of the immune system, autoimmune diseases, hepatic and renal
disease, lymphomas, inflammation, allergies, ischemic shock, Alzheimer's
and cognitive disorders, schizophrenia, prostate diseases, obesity,
disorders involving osteoclasts such as osteoporosis, arthritis or
malignancies, diseases of testes, lung or thymus, digestive/endocrine
disorders, infections and AIDS. The polypeptides are also useful for
identifying their binding partners.
XX
SQ Sequence 610 BP; 187 A; 121 C; 103 G; 196 T; 3 other;

Query Match      13.0%; Score 30; DB 20; Length 610;
Best Local Similarity 60.0%; Pred. No. 3;
Matches 48; Conservative 1; Mismatches 31; Indels 0; Gaps 0;

QY 92 AGTTCCCTTCCCCAGGAAACACACGAGCTGCGCAAGACCTCTCTGGTGATCGAG 151
    |||||
Db 258 AGTTCCCTTCCCCAGGAAACACACGAGCTGCGCAAGACCTCTCTGGTGATCGAG 151
    |||||

QY 152 CCTAAGGGATGTTTGTGT 171
    |||||
Db 198 CCAAGGTAATTTTGTGT 179
    |||||

RESULT 15
AAS32249
ID AAS32249 standard; DNA; 32187 BP.
XX
AC AAS32249;
XX
XX
XX 04-DEC-2001 (first entry)
XX
XX Human DNA repair and processing genomic DNA #35.
XX
XX DNA processing; human; mouse; rabbit; goat; horse; cat; gene therapy;
XX chicken; sheep; immunosuppressive; antiarthritic; vasotropic;
XX antirheumatic; antiproliferative; cytostatic; cardiac; neuroprotective;
XX cerebroprotective; vulnertic; gene therapy; autoimmune disease; cancer;
XX ophthalmological; liver; cardiovascular disorder; ds;
XX hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;
XX cerebrovascular disorder; nervous system disorder; bacterial infection;
XX fungal infection; viral infection; ocular disorder; endocrine disorder;
XX gastrointestinal disorder; renal disorder; respiratory disorder; dog;
XX wound healing; skin aging; organ transplantation; tissue regeneration;
XX anti-infertility; DNA repair protein.
XX
XX Homo sapiens.
XX
XX WO200155204-A1.
XX
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01336.
XX
XX 31-JAN-2000; 2000US-0179065.
XX
XX 04-FEB-2000; 2000US-0180528.
XX
XX 24-FEB-2000; 2000US-0184664.
XX
XX 02-MAR-2000; 2000US-0186350.
XX
XX 16-MAR-2000; 2000US-0189874.
XX
XX 17-MAR-2000; 2000US-0190076.
XX
XX 18-APR-2000; 2000US-0198123.
XX
XX 19-MAY-2000; 2000US-0205515.
XX
XX 07-JUN-2000; 2000US-0209467.
XX
XX 28-JUN-2000; 2000US-0214886.
XX
XX 30-JUN-2000; 2000US-0215135.
XX
XX 07-JUL-2000; 2000US-0216647.
XX
XX 07-JUL-2000; 2000US-0216880.
XX
XX 11-JUL-2000; 2000US-0217487.

```

PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220863.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 18-AUG-2000; 2000US-0225759.
PR 22-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232081.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.

PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249267.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 05-JAN-2001; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-496846/54.

Nucleic acids encoding human polypeptides, useful for preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease and diabetic retinopathy -

Claim 1; SEQ ID No 109; 460pp; English.

Sequences AAS32215-AAS32250 represent genomic DNA molecules, which encode the DNA repair and processing polypeptides of the invention. DNA repair and processing polypeptides and their associated polynucleotides are useful in the diagnosis, treatment and prevention of various types of disorders in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological condition can be determined by detecting the presence or absence of a mutation in a DNA repair and processing polynucleotide. The treatable disorders include autoimmune diseases such as rheumatoid arthritis, hyperproliferative disorders such as neoplasms of the breast or liver, cardiovascular disorders such as cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,

CC nervous system disorders such as Alzheimer's disease, infections caused
CC by bacteria, viruses and fungi, ocular disorders such as corneal
CC infection, endocrine disorders such as premature labour and infertility,
CC gastrointestinal disorders such as Crohn's disease, renal disorders such
CC as glomerulonephritis and respiratory disorders such as asthma and
CC pleurisy. The polypeptides can also be used to aid wound healing, to
CC prevent skin aging due to sunburn, to maintain organs before
CC transplantation, to regenerate tissues and in chemotaxis.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

```
Query Match      12.7%; Score 29.4; DB 22; Length 32187;
Best Local Similarity 49.1%; Pred. No. 20;
Matches 78; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Qy 37 TTGCTGTGGCATCCTGTGGAGGCACGCTCTGATTCATGATGAGGTTTCAGTGTCTCTAGTTTC 96
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20285 TTGCAGGTGGGGCCCTAGTGGGAGGTGCTGAGTTATAGGGATGGCTTCGTGAATGGCTTC 20344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 97 CCTTCCCCCAGGAAAAACGACACGGGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTAA 156
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20345 GTGCCCTTCTGGCCGTAATGAGCGGGTGTGTTAAACACACTGGTGTGTTTAAAGAGTGTA 20404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Qy 157 GGGATGGTGTGTTAGGGCCCTATGCTTGCACACTGG 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 20405 GATCTGTTAAGAGTGTGGCACCTTTCCTCTCTCTTG 20443
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 24, 2002, 12:37:23
Job time : 300 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 12:30:11 ; Search time 1978 Seconds
(without alignments)
1891.384 Million cell updates/sec

Title: US-09-673-716-1
Perfect score: 231
Sequence: 1 gtgggggcaggctaagca.....ttcaccatgagggtgctt 231

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pin:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	231	100.0	477	17	AZ231159
2	231	100.0	501	17	AZ440299
3	231	100.0	527	12	BG087407
c 4	231	100.0	558	17	BH036520
5	231	100.0	559	17	BH105759
c 6	231	100.0	577	17	AZ420721

c 7	231	100.0	587	17	AZ832364
c 8	231	100.0	627	17	AZ495883
c 9	231	100.0	647	17	AZ993990
c 10	231	100.0	721	17	AZ994818
c 11	229.4	99.3	472	17	AZ310386
c 12	229.4	99.3	505	17	AZ054649
c 13	229.4	99.3	518	10	AZ558986
c 14	229.4	99.3	534	17	AZ341190
c 15	229.4	99.3	535	17	AZ326140
c 16	229.4	99.3	539	17	AZ965228
c 17	229.4	99.3	555	13	BH219777
c 18	229.4	99.3	558	17	AZ987016
c 19	229.4	99.3	584	17	AZ432413
c 20	229.4	99.3	588	17	AZ242504
c 21	229.4	99.3	631	10	AZ554086
c 22	229.4	99.3	662	17	AZ333042
c 23	229.4	99.3	686	17	BH092157
c 24	229.4	99.3	707	17	BH117452
c 25	229.4	99.3	834	17	BH062187
c 26	228.4	98.9	606	13	BH134544
c 27	227.8	98.6	309	17	BH068706
c 28	227.8	98.6	457	17	AZ071850
c 29	227.8	98.6	492	10	AZ552509
c 30	227.8	98.6	531	17	AZ636367
c 31	227.8	98.6	557	17	BH078165
c 32	227.8	98.6	559	14	BQ553681
c 33	227.8	98.6	596	17	AZ830822
c 34	227.8	98.6	646	17	BH099567
c 35	227.8	98.6	678	17	AZ897416
c 36	227.8	98.6	686	17	AZ259805
c 37	227.8	98.6	706	17	AZ797045
c 38	227.8	98.6	793	17	AZ719135
c 39	227.8	98.6	838	17	BH027115
c 40	227.4	98.4	553	17	AZ792711
c 41	227.4	98.4	642	17	AZ090071
c 42	227.4	98.4	675	17	AZ957637
c 43	226.8	98.2	471	14	BH877325
c 44	226.8	98.2	570	17	AQ970072
c 45	226.8	98.2	641	17	AZ596677

ALIGNMENTS

RESULT 1
AZ231159
LOCUS RPCI-23-51K23.TJ RPCI-23 Mus musculus genomic clone RPCI-23-51K23, 477 bp DNA linear GSS 14-JUN-2000
DEFINITION DNA sequence.
ACCESSION AZ231159
VERSION AZ231159.1 GI:8539205
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 477)
AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S., Akinret, B., Levins, M., McGann, S., Tsegaye, G., Geer, K., Krol, M., de Jong, P. and Fraser, C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Unpublished (1999)
COMMENT Other_GSSs: RPCI-23-51K23.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
 or from Resea ch Genetics (inforesgen.com). BAC end page:
http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 51 row: K column: 23
 Seq primer: SP6
 Class: BAC ends.

FEATURES

source
 Location/Qualifiers
 1. .477
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-23-51K23"
 /clone_lib="RPCI-23"
 /sex="Female"
 /lab_host="DH10B"
 /note="Organ: Kidney/Brain; Vector: pBACE3.6; Site1:
 EcoRI; Site2: EcoRI; Female C57BL/6J mouse kidney and/or
 brain genomic DNA was isolated and partially digested
 with a combination of EcoRI and EcoRI Methyase. Size
 selected DNA was cloned into the pBACE3.6 vector at the
 EcoRI sites. The ligation products were transformed into
 DH10B electrocompetent cells (BRL Life Technologies)."

BASE COUNT 94 a 110 c 143 g 130 t
 ORIGIN
 Query Match 100.0%; Score 231; DB 17; Length 477;
 Best Local Similarity 100.0%; Pred. No. 7.2e-61;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGGAGGCTAAGCACTGCACAGGATAGCTTGTGTGCATCTGTGGAAGG 60
 Db 232 GTGGGTGGAGGCTAAGCACTGCACAGGATAGCTTGTGTGCATCTGTGGAAGG 291
 QY 61 CAGCTCTGATTGTCATGAAGTTTCAGTGTCTAGTTCCTTCCCTCCAGGAAAAACGACACG 120
 Db 292 CAGCTCTGATTGTCATGAAGTTTCAGTGTCTAGTTCCTTCCCTCCAGGAAAAACGACACG 351
 QY 121 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTGTAGGGCCCT 180
 Db 352 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTGTAGGGCCCT 411
 QY 181 ATGCTTGCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 231
 Db 412 ATGCTTGCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 462

RESULT 2
 AZ440299 501 bp DNA linear GSS 03-Oct-2000
 LOCUS 1M0231K17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M0231K17 F, DNA sequence.
 ACCESSION AZ440299
 VERSION AZ440299.1 GI:10564312
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 501)
 AUTHORS Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C.,
 Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
 M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
 and Wright,D., Weiss,R.

TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts

JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177

Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0231 row: K column: 17
 Seq primer: CGTTGTAAACGACGCGCCACT
 Class: plasmid ends
 High quality sequence stop: 501.
 Location/Qualifiers
 1. .501

FEATURES

source
 Location/Qualifiers
 1. .501
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0231K17"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male)' was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/dnares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (gi14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT 119 a 123 c 136 g 123 t
 ORIGIN
 Query Match 100.0%; Score 231; DB 17; Length 501;
 Best Local Similarity 100.0%; Pred. No. 7.4e-61;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGGAGGCTAAGCACTGCACAGGATAGCTTGTGTGCATCTGTGGAAGG 60
 Db 94 GTGGGTGGAGGCTAAGCACTGCACAGGATAGCTTGTGTGCATCTGTGGAAGG 153
 QY 61 CAGCTCTGATTGTCATGAAGTTTCAGTGTCTAGTTCCTTCCCTCCAGGAAAAACGACACG 120
 Db 154 CAGCTCTGATTGTCATGAAGTTTCAGTGTCTAGTTCCTTCCCTCCAGGAAAAACGACACG 213
 QY 121 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTGTAGGGCCCT 180
 Db 214 GGAGCTGGCCAAAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTGTAGGGCCCT 273
 QY 181 ATGCTTGCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 231
 Db 274 ATGCTTGCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 324

RESULT 3

BG087407 527 bp mRNA linear EST 26-JAN-2001
 H3139B12-5 NIA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3139B12 5', mRNA sequence.

ACCESSION BG087407
 VERSION BG087407.1 GI:12569971
 KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 527)
 AUTHORS Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka
 T.S., Carter,M.G. and Ko,M.S.H.

TITLE Verification and initial annotation of NIA mouse 15K cDNA clone set
JOURNAL Unpublished (2001)
COMMENT Other_ESTS: H3139B12-3
Contact: George J. Kargul
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cda@lgsun.grc.nia.nih.gov
This clone set has been freely distributed to the community. Please
visit <http://lgsun.grc.nia.nih.gov/cDNA/15k.html> for details.
Plate: H3139 row: B column: 12
Seq primer: -21M13 Reverse
High quality sequence stop: 527
POLYA-No.

FEATURES

source

1. .527
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="niaEST:H3139B12-5"
/db_xref="taxon:10090"
/clone="H3139B12"
/dev_stage="Clones arrayed from a variety of cDNA libraries"
/lab_host="DH10B"
/note="Vector: pSPOR1; Site_1: SalI; Site_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryo cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos, and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 kb. All source libraries are cloned unidirectionally with Oligo(dt)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U.S.A., 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."

BASE COUNT 104 a 121 c 158 g 144 t
ORIGIN

Query Match 100.0%; Score 231; DB 12; Length 527;
Best Local Similarity 100.0%; Pred. No. 7.6e-61;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGGCGAGCTAAGCACTGCACAGAGGATAGCTTGTGTTGGCATCCTGTGGAAGG 60
|||||
Db 236 GTGGGTGGCGAGCTAAGCACTGCACAGAGGATAGCTTGTGTTGGCATCCTGTGGAAGG 295
|||||
QY 61 CAGCTCTGATTCATGAAGTTCAAGTCTAGTCTCCCTTCCCCAGGAAAAACGACACG 120
|||||
Db 296 CAGCTCTGATTCATGAAGTTCAAGTCTAGTCTCCCTTCCCCAGGAAAAACGACACG 355
|||||
QY 121 GGAGCTGGCCAGACCTCTCTGGGTATGAGCTTAGGATGTTTGTAGGGCCCCCT 180
|||||
Db 356 GGAGCTGGCCAGACCTCTCTGGGTATGAGCTTAGGATGTTTGTAGGGCCCCCT 415
|||||
QY 181 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCCATGAGGCTTGCTT 231
|||||
Db 416 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCCATGAGGCTTGCTT 466
|||||

RESULT 4

BH036520/c

LOCUS

DEFINITION RPCI-24-254C6.TJ RPCI-24 Mus musculus genomic clone RPCI-24-254C6,

558 bp

DNA

linear

GSS 17-JUL-2001

DNA sequence.

BH036520

BH036520.1 GI:14811469

GSS.

house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 558)

Zhao,S., Nierman,W., Malek,J., Shatsman,S., Akinret,B., Levins,M.,
Tsegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E.,
Russell,D., de Jong,P. and Fraser,C.M.

Mouse BAC End Sequences from Library RPCI-24

Unpublished (1999)

Other_GSSs: RPCI-24-254C6.TV

Contact: Shaying Zhao

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: szhao@tigr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong

(pdejong@mail.cho.org). Clones may be purchased from BACPAC

Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end

page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html

Plate: 254 row: C column: 6

Seq primer: SP6

Class: BAC ends.

Location/Qualifiers

1. .558

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="RPCI-24-254C6"

/clone_lib="RPCI-24"

/sex="Male"

/cell_type="Spleen/Brain"

/notes="Vector: pTARBAC1; Site_1: BamHI; Site_2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the
BamHI sites using MboI partially digested male C57BL/6J
DNA."

BASE COUNT 151 a 158 c 129 g 120 t
ORIGIN

Query Match 100.0%; Score 231; DB 17; Length 558;
Best Local Similarity 100.0%; Pred. No. 7.8e-61;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTGGCGAGCTAAGCACTGCACAGAGGATAGCTTGTGTTGGCATCCTGTGGAAGG 60
|||||
Db 376 GTGGGTGGCGAGCTAAGCACTGCACAGAGGATAGCTTGTGTTGGCATCCTGTGGAAGG 317
|||||
QY 61 CAGCTCTGATTCATGAAGTTCAAGTCTAGTCTCCCTTCCCCAGGAAAAACGACACG 120
|||||
Db 316 CAGCTCTGATTCATGAAGTTCAAGTCTAGTCTCCCTTCCCCAGGAAAAACGACACG 257
|||||
QY 121 GGAGCTGGCCAGACCTCTCTGGGTATGAGCTTAGGATGTTTGTAGGGCCCCCT 180
|||||
Db 256 GGAGCTGGCCAGACCTCTCTGGGTATGAGCTTAGGATGTTTGTAGGGCCCCCT 197
|||||
QY 181 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCCATGAGGCTTGCTT 231
|||||
Db 196 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCCATGAGGCTTGCTT 146
|||||

RESULT 5

BH105759

LOCUS

DEFINITION RPCI-24-237C16.TJ RPCI-24 Mus musculus genomic clone RPCI-24-237C16,

559 bp

DNA

linear

GSS 19-JUL-2001

ACCESSION BH105759
 VERSION BH105759.1 GI:14935045
 KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 559)
 AUTHORS Zhao, S., Niernan, W., Malek, J., Shatsman, S., Akinret, B., Levins, M., Tsengaye, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregeorgis, E., Russell, D., de Jong, P., and Fraser, C. M.
 TITLE Mouse BAC End Sequences from Library RPCI-24
 JOURNAL Unpublished (1999)
 COMMENT Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org
 Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (<http://www.choi.org/bacpac/orderingframe.htm>). BAC end page: http://www.tigr.org/tdb/bac_ends/mouse/bac_end_intro.html
 Plate: 237 row: C column: 16
 Seq primer: SP6
 Class: BAC ends.
 Location/Qualifiers
 1. 559
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="RPCI-24-237C16"
 /clone_lib="RPCI-24"
 /sex="Male"
 /cell_type="Spleen/Brain"
 /note="Vector: pRABAC1; Site_1: BamH1; Site_2: BamH1; RPCI-24 Mouse BAC Library produced by Pieter de Jong. The library was cloned in the pRABAC1 cloning vector at the BamH1 sites using MboI partially digested male C57BL/6J DNA."
 BASE COUNT 119 a 132 c 159 g 149 t
 ORIGIN
 Query Match 100.0%; Score 231; DB 17; Length 559;
 Best Local Similarity 100.0%; Pred. No. 7.8e-61;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGGGTGCGAGGCTAGCACTGCACAGAGATAGCTTGTGGCATCTGTGGAGG 60
 DB 243 GTGGGTGCGAGGCTAGCACTGCACAGAGATAGCTTGTGGCATCTGTGGAGG 302
 QY 61 CACGCTCTGATTGTCATGAAGTTCAGTCTCCTAGTTCCTTCCCCAGAAAACGACACG 120
 DB 303 CACGCTCTGATTGTCATGAAGTTCAGTCTCCTAGTTCCTTCCCCAGAAAACGACACG 362
 QY 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGTTTGTGTAGGGCCCT 180
 DB 363 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGTTTGTGTAGGGCCCT 422
 QY 181 ATGCTTCACACTGGGGATCAGACCTCTACCTTCACCCATGAGCTTGCTT 231
 DB 423 ATGCTTCACACTGGGGATCAGACCTCTACCTTCACCCATGAGCTTGCTT 473
 RESULT 6
 AZ420721/c 577 bp DNA linear GSS 03-OCT-2000
 LOCUS IM0198L07R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
 DEFINITION clone UUGC1M0198L07 R, DNA sequence.
 ACCESSION AZ420721
 VERSION AZ420721.1 GI:10544734

KEYWORDS GSS.
 SOURCE house mouse.
 ORGANISM Mus musculus
 REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 577)
 AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamill, C., Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A., and Wright, D., Weiss, R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
 JOURNAL Unpublished (2000)
 COMMENT Contact: Robert B. Weiss
 University of Utah Genome Center
 Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: ddunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0198 row: L column: 07
 Seq primer: CACACAGGAACACGCTATGACC
 Class: plasmid ends
 High quality sequence stop: 577.
 Location/Qualifiers
 1. 577
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGC1M0198L07"
 /clone_lib="Mouse 10kb plasmid UUGC1M library"
 /sex="Male"
 /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from the Jackson Laboratory Mouse DNA Resource (<http://www.jax.org/resources/documents/dnares/>). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gblAF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
 BASE COUNT 152 a 171 c 131 g 123 t
 ORIGIN
 Query Match 100.0%; Score 231; DB 17; Length 577;
 Best Local Similarity 100.0%; Pred. No. 8e-61;
 Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GTGGGTGCGAGGCTAAGCACTGCACAGAGATAGCTTGTGGCATCTGTGGAGG 60
 DB 374 GTGGGTGCGAGGCTAAGCACTGCACAGAGATAGCTTGTGGCATCTGTGGAGG 315
 QY 61 CACGCTCTGATTGTCATGAAGTTCAGTCTCCTAGTTCCTTCCCCAGAAAACGACACG 120
 DB 314 CACGCTCTGATTGTCATGAAGTTCAGTCTCCTAGTTCCTTCCCCAGAAAACGACACG 255
 QY 121 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGTTTGTGTAGGGCCCT 180
 DB 254 GGAGCTGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGGATGTTTGTGTAGGGCCCT 195

```

QY 181 ATGCTTGACACTGGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 231
Db 194 ATGCTTGACACTGGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 144

RESULT 7
AZ832364/c
LOCUS 2M0112B13R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
DEFINITION clone UUGC2M0112B13 R, DNA sequence.
ACCESSION AZ832364
VERSION AZ832364.1 GI:13002272
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 587)
Islam,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0112 row: B column: 13
Seq primer: CACACAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 587.
FEATURES
source
1..587
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="UUGC2M0112B13"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/sex="Male"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 159 a 168 c 135 g 125 t
ORIGIN

QY 1 GTGGGTGCGAGCTAAGCACTGCACAGAGATAGCTTGCTGTGGCATCCTGTGGAAG 60
Db 246 GTGGGTGCGAGCTAAGCACTGCACAGAGATAGCTTGCTGTGGCATCCTGTGGAAG 187

QY 61 CACGCTCTGATTGCATGAAGGTTTCAGTGTCTCCTTCCTCCCTCCAGGAAAAAGACACG 120
Db 186 CACGCTCTGATTGCATGAAGGTTTCAGTGTCTCCTTCCTCCCTCCAGGAAAAAGACACG 127

QY 121 GGAGCTGGCCCAACCTCTCTGGTGATGACCTAAGGATGGTTTGTGTAGGGCCCT 180
Db 126 GGAGCTGGCCCAACCTCTCTGGTGATGACCTAAGGATGGTTTGTGTAGGGCCCT 67

QY 181 ATGCTTGACACTGGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 231
Db 66 ATGCTTGACACTGGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCTT 16

RESULT 8
AZ495883/c
LOCUS 1M0332E03F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
DEFINITION clone UUGC1M0332E03 F, DNA sequence.
ACCESSION AZ495883
VERSION AZ495883.1 GI:10671633
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 627)
Islam,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
TITLE Mouse whole genome scaffolding with paired end reads from 10kb
JOURNAL plasmid inserts
COMMENT Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0332 row: E column: 03
Seq primer: CGTTGTAACAGACGCGCCAGT
Class: plasmid ends
High quality sequence stop: 627.
FEATURES
source
1..627
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone_lib="UUGC1M0332E03"
/lab_host="E. Coli strain XL10-Gold, TI-resistant, F-"
/sex="Male"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
BASE COUNT 159 a 168 c 135 g 125 t
ORIGIN

```

Query Match
Best Local Similarity 100.0%; Score 231; DB 17; Length 587;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 160 a 190 c 142 g 135 t
ORIGIN

Query Match 100.0%; Score 231; DB 17; Length 627;
Best Local Similarity 100.0%; Pred. No. 8.4e-61;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTCGCGAGGCTAGCACTGCACAGAGATAGCTTGTGTGGCATCTCTGTGGAAGG 60

Db 441 GTGGGTCGCGAGGCTAGCACTGCACAGAGATAGCTTGTGTGGCATCTCTGTGGAAGG 382

QY 61 CACGCTCTGATTGTCATGAAGTTTCAGTCTCCTTCCCTCCCGAGGAAACGACACG 120

Db 381 CACGCTCTGATTGTCATGAAGTTTCAGTCTCCTTCCCTCCCGAGGAAACGACACG 322

QY 121 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGCCCT 180

Db 321 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGCCCT 262

QY 181 ATGCTTCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCCTT 231

Db 261 ATGCTTCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCCTT 211

RESULT 9

LOCUS

DEFINITION 2M0279D16F Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0279D16 F, DNA sequence.

ACCESSION

AZ993990

VERSION

AZ993990.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 647)

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000

Std Error: 0.00

Plate: 0279

row: D

column: 16

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

High quality sequence stop: 647.

FEATURES

source

1. .647

Location/Qualifiers

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUGC2M0279D16"

/clone_lib="Mouse 10kb plasmid UUGC2M library"

/sex="Female"

/lab_host="E. coli strain XL10-Gold, TI-resistant, F-"

/note="Vector: pWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson

laboratory mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adapted DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gil4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 149 a 146 c 185 g 167 t
ORIGIN

Query Match 100.0%; Score 231; DB 17; Length 647;
Best Local Similarity 100.0%; Pred. No. 8.5e-61;
Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGGGTCGCGAGGCTAGCACTGCACAGAGATAGCTTGTGTGGCATCTCTGTGGAAGG 60

Db 235 GTGGGTCGCGAGGCTAGCACTGCACAGAGATAGCTTGTGTGGCATCTCTGTGGAAGG 294

QY 61 CACGCTCTGATTGTCATGAAGTTTCAGTCTCCTTCCCTCCCGAGGAAACGACACG 120

Db 295 CACGCTCTGATTGTCATGAAGTTTCAGTCTCCTTCCCTCCCGAGGAAACGACACG 354

QY 121 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGCCCT 180

Db 355 GGAGCTGGCCCAAGACCTCTCTGGGTGATGAGCCTTAAGGATGTTTGTGTAGGCCCT 414

QY 181 ATGCTTCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCCTT 231

Db 415 ATGCTTCACACTGGGATCAGACCTCTACCTTCACCCATGAGGCTTGCCTT 465

RESULT 10

LOCUS

DEFINITION 2M0280J24F Mouse 10kb plasmid UUGC2M library Mus musculus genomic clone UUGC2M0280J24 F, DNA sequence.

ACCESSION

AZ994818

VERSION

AZ994818.1

KEYWORDS

GSS.

SOURCE

house mouse.

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 721)

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

JOURNAL

Unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah

University of Utah

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunn@genetics.utah.edu

Insert Length: 10000

Std Error: 0.00

Plate: 0280

row: J

column: 24

Seq primer: CGTTGTAAACGACGCCAGT

Class: plasmid ends

```
FEATURES
  source
    High quality sequence stop: 721.
    Location/Qualifiers
      1..721
        /organism="Mus musculus"
        /strain="C57BL/6J"
        /db_xref="taxon:10090"
        /clone="UUC2M0280324"
        /clone_lib="Mouse 10kb plasmid UUC2M library"
        /sex="Female"
        /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
        /note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (female) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
      BASE COUNT 191 a 202 c 166 g 162 t
      ORIGIN
        Query Match 100.0%; Score 231; DB 17; Length 721;
        Best Local Similarity 100.0%; Pred. No. 9.1e-61;
        Matches 231; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
        QY 1 GTGGGTGCGAGCTAAGCACTGCACAGAGTAGCTGCTGTGGCATCCTGTGGAAGG 60
        Db 300 GTGGGTGCGAGCTAAGCACTGCACAGAGTAGCTGCTGTGGCATCCTGTGGAAGG 241
        QY 61 CAGCTCTGATTCATGAAGGTTTCAGTGTCTCTATGTTCCCTTCCCCAGAGAAACGACACG 120
        Db 240 CAGCTCTGATTCATGAAGGTTTCAGTGTCTCTATGTTCCCTTCCCCAGAGAAACGACACG 181
        QY 121 GGAGCTGGCCAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTAGGGCCCT 180
        Db 180 GGAGCTGGCCAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTAGGGCCCT 121
        QY 181 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCATGAGCTTGCTT 231
        Db 120 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCATGAGCTTGCTT 70
      RESULT 11
      LOCUS AZ310386/c
      DEFINITION 472 bp DNA linear GSS 29-SEP-2000
      ACCESSION 1M0025013F Mouse 10kb plasmid UUC1M library Mus musculus genomic
      VERSION clone UUC1M0025013 F, DNA sequence.
      KEYWORDS AZ310386
      SOURCE GSS.
      ORGANISM
        house mouse.
        Mus musculus
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
      REFERENCE
        1 (bases 1 to 472)
      AUTHORS
        Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
        Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
        ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
        and Wright,D.,Weiss,R.
      TITLE Mouse whole genome scaffolding with paired end reads from 10kb
        plasmid inserts
      JOURNAL Unpublished (2000)
```

COMMENT

Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0025 row: O column: 13
Seq primer: CGTTGTAACAGCAGCGCCAGT
Class: plasmid ends
High quality sequence stop: 472.
Location/Qualifiers
1..472
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUC1M0025013"
/clone_lib="Mouse 10kb plasmid UUC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 133 a 140 c 104 g 95 t
ORIGIN

Query Match 99.3%; Score 229.4; DB 17; Length 472;
Best Local Similarity 99.6%; Pred. No. 2.2e-60;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GTGGGTGCGAGCTAAGCACTGCACAGAGTAGCTGCTGTGGCATCCTGTGGAAGG 60
Db 281 GTGGGTGCGAGCTAAGCACTGCACAGAGTAGCTGCTGTGGCATCCTGTGGAAGG 222

QY 61 CAGCTCTGATTCATGAAGGTTTCAGTGTCTCTATGTTCCCTTCCCCAGAGAAACGACACG 120
Db 221 CAGCTCTGATTCATGAAGGTTTCAGTGTCTCTATGTTCCCTTCCCCAGAGAAACGACACG 162

QY 121 GGAGCTGGCCAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTAGGGCCCT 180
Db 161 GGAGCTGGCCAGACCTCTCTGGGTGATGAGCCTAAGGATGTTTGTAGGGCCCT 102

QY 181 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCATGAGCTTGCTT 231
Db 101 ATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCCATGAGCTTGCTT 51

RESULT 12
LOCUS AZ054649
DEFINITION RPCI-23-401C1 TJ RPCI-23 Mus musculus genomic clone RPCI-23-401C1,
DNA sequence.
ACCESSION AZ054649
VERSION AZ054649.1 GI:7345885
KEYWORDS GSS.

Query Match 99.3%; Score 229.4; DB 17; Length 534;
Best Local Similarity 99.6%; Pred. No. 2.4e-60;
Matches 230; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 24, 2002, 12:31:11 ; Search time 48 Seconds
(without alignments)
1475.882 Million cell updates/sec

Title: US-09-673-716-1
Perfect score: 231
Sequence: 1 gtggggtcggagctaagca.....ttcaccatgaggttgctt 231

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents.NA.*
1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2.6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2.6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2.6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86.2	37.3	1434	3	US-08-695-191-13
2	86.2	37.3	1434	3	US-08-682-080-13
3	81.4	35.2	1400	3	US-08-695-191-14
4	81.4	35.2	1400	3	US-08-682-080-14
5	29.2	12.6	540	4	US-09-250-609-95
c 6	28.6	12.4	2692	1	US-07-932-454A-2
7	27.6	11.9	1665	3	US-08-766-354A-29
c 8	27.2	11.8	2297	2	US-08-588-983-21
c 9	27.2	11.8	2297	2	US-08-588-976-21
10	27	11.7	1054	6	5189147-4
11	26.8	11.6	503	4	US-09-250-609-85
12	26.8	11.6	3213	2	US-08-633-770A-4
13	26.4	11.4	13865	3	US-09-009-217-11
14	26.4	11.4	13865	3	US-09-009-656-11
15	26.2	11.3	71	1	US-08-434-001-210
16	26.2	11.3	71	1	US-08-433-585-210
17	26.2	11.3	71	1	US-08-434-425-210
18	26.2	11.3	71	2	US-08-437-667-210
19	26.2	11.3	71	3	US-08-906-955-210
20	26.2	11.3	71	3	US-08-945-909-210
21	26.2	11.3	71	4	US-09-396-002A-210
22	26.2	11.3	71	5	PCI-US96-06060-210
c 23	25.8	11.2	444	4	US-08-928-213B-180
24	25.8	11.2	5865	4	US-09-011-745-8
25	25.8	11.2	9058	4	US-08-913-014A-9
c 26	25.6	11.1	419	4	US-09-214-095D-113
27	25.6	11.1	4279	4	US-09-041-886-22

c 28	25.2	10.9	371	1	US-08-620-467A-1	Sequence 1, Appl
c 29	25.2	10.9	371	1	US-08-348-572-1	Sequence 1, Appl
c 30	25.2	10.9	371	3	US-09-041-090B-1	Sequence 1, Appl
31	25.2	10.9	2336	4	US-09-228-986-10	Sequence 10, Appl
32	25	10.8	385	4	US-08-905-223-208	Sequence 208, App
33	25	10.8	1232	4	US-09-392-184-4	Sequence 4, Appl
c 34	25	10.8	4020	4	US-09-050-159-130	Sequence 130, App
c 35	25	10.8	4024	4	US-09-162-484-18	Sequence 18, Appl
c 36	25	10.8	4550	4	US-09-462-136-1	Sequence 1, Appl
c 37	25	10.8	8323	1	US-08-110-300A-8	Sequence 8, Appl
c 38	25	10.8	8323	2	US-08-886-642-8	Sequence 2, Appl
c 39	25	10.8	8323	4	US-09-433-322B-2	Sequence 2, Appl
c 40	25	10.8	8323	5	PCT-US93-08041-8	Sequence 8, Appl
41	25	10.8	10367	1	US-08-110-300A-9	Sequence 9, Appl
42	25	10.8	10367	2	US-08-886-642-9	Sequence 9, Appl
43	25	10.7	10367	5	PCT-US93-08041-9	Sequence 9, Appl
c 44	24.8	10.7	7301	4	US-09-816-088-3	Sequence 3, Appl
45	24.6	10.6	361	4	US-09-025-769B-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-08-695-191-13
; Sequence 13, Application US/08695191
; Patent No. 6025155
; GENERAL INFORMATION:
; APPLICANT: Hadlaczky, Gyula
; APPLICANT: Szalay, Aladar
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND
; TITLE OF INVENTION: METHODS FOR PREPARING ARTIFICIAL CHROMOSOMES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/695,191
; FILING DATE: 07-AUG-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 15-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/629,822
; FILING DATE: 10-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6869-402C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1434 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO


```

Best Local Similarity 71.9%; Pred. No. 4.6e-20;
Matches 123; Conservative 0; Mismatches 41; Indels 7; Gaps 1;

QY 55 GGAAGGCACGCTCTGATTGCATGAAGGTTCAAGTGCTCCTAGTTCCTTCCCTCCCGCAGGAAAAAC 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 771 GGGAGCATGTCATCTCTTCAAGAAGGTTGAGTGCTCAAGTGCTCTTCTCCAGCGAAAAAC 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 115 GACACGGGAGCTGCCAAGACCTCTCTGGGTGA-----TGAGCCTTAAGGATGGTTTT 167
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 831 GACACGGGAGCAGTCAAGGTTGCTCTGGGTAAAAAGCCTGTGAGCCTAAGAGCTAATCCT 890
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 168 GTGTAGGGCCCCATGCTTGCACACTGGGGATCAGACCTCTACCTTCACCC 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 891 GTACATGGCTCTTTACTACACACTGGGATTTGACCTCTATCTCCACTC 941
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-08-682-080-14
; Sequence 14, Application US/08692080
; Patent No. 6077697
; GENERAL INFORMATION:
; APPLICANT: Hadlaczky, Gyula
; APPLICANT: Szalay, Aladar
; TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND METHODS FOR
; TITLE OF INVENTION: PREPARING ARTIFICIAL CHROMOSOMES
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,080
; FILING DATE: 10-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/629,822
; FILING DATE: 10-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6869-402B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-238-0999
; TELEFAX: 619-238-0062
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1400 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
;
US-08-682-080-14

Query Match 35.2%; Score 81.4; DB 3; Length 1400;
Best Local Similarity 71.9%; Pred. No. 4.6e-20;
Matches 123; Conservative 0; Mismatches 41; Indels 7; Gaps 1;

QY 55 GGAAGGCACGCTCTGATTGCATGAAGGTTCAAGTGCTCCTAGTTCCTTCCCTCCCGCAGGAAAAAC 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 771 GGGAGATGTCATCTTTTCAAGAAGGTTTGAGTGTCCAAAGTGTCTCTCCAGGCAAAAC 830
QY 115 GACACGGGAGCTGGCCAAGACCTCTCTGGGTGA-----TCAGCCTAAGGGATGGTTTT 167
Db 831 GACACGGGAGCAGGTGAGGTGCTCTGGGTAAAGCCCTGTGAGCCTAAAGACTAATCCT 890
QY 168 GTGTAGGGCCCCATGCTTTCACACATGGGGATGACAGACCTCTTACCTTCACCC 218
Db 891 GTACATGGCTCCTTTTACCTACACATGGGGATTTGACCTCTTATCTCCACTC 941

RESULT 5
US-09-250-609-95
; Sequence 95, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-95

Query Match 12.6%; Score 29.2; DB 4; Length 540;
Best Local Similarity 54.7%; Pred. No. 0.36;
Matches 58; Conservative 0; Mismatches 48; Indels 0; Gaps

QY 26 CACAGGATAGCTTGCTTGGCATCTGTGGAGGACACCTCTGATGTCATGAAGTTTCAG 85
Db 227 CAGAACCTGTCTAGGAGCTGGCATGATGTGCAGGTCTCTACTGCTACAAGAAGACTCAA 286
QY 86 TGTCTTAGTTCCCTTCCCCAGAAACGACACGGAGCTGGCCA 131
Db 287 GAAACTCTTTCACAGGCTGGACAGAAACATCATGCTGCCCTGTCCA 332

RESULT 6
US-07-932-454A-2/c
; Sequence 2, Application US/07932454A
; Patent No. 5262318
; GENERAL INFORMATION:
; APPLICANT: GUTHRIE, ELLEN P.
; TITLE OF INVENTION: ISOLATED DNA ENCODING THE SphI
; TITLE OF INVENTION: RESTRICTION ENDONUCLEASE AND RELATED METHODS FOR
; TITLE OF INVENTION: THE SAME
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/932,454A
; FILING DATE: 19920820
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: 42078

```

```

Db 771 GGGAGATGTCATCTTTTCAAGAAGGTTTGAGTGTCCAAAGTGTCTCTCCAGGCAAAAC 830
QY 115 GACACGGGAGCTGGCCAAGACCTCTCTGGGTGA-----TCAGCCTAAGGGATGGTTTT 167
Db 831 GACACGGGAGCAGGTGAGGTGCTCTGGGTAAAGCCCTGTGAGCCTAAAGACTAATCCT 890
QY 168 GTGTAGGGCCCCATGCTTTCACACATGGGGATGACAGACCTCTTACCTTCACCC 218
Db 891 GTACATGGCTCCTTTTACCTACACATGGGGATTTGACCTCTTATCTCCACTC 941

RESULT 5
US-09-250-609-95
; Sequence 95, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250,609A
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 95
; LENGTH: 540
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-250-609-95

Query Match 12.6%; Score 29.2; DB 4; Length 540;
Best Local Similarity 54.7%; Pred. No. 0.36;
Matches 58; Conservative 0; Mismatches 48; Indels 0; Gaps

QY 26 CACAGGATAGCTTGCTTGGCATCTGTGGAGGACACCTCTGATGATGAAGTTTCAG 85
Db 227 CAGAACCTGTCTAGGAGCTGGCATGATGTGCAGGTCTCTACTGCTACAAGAAGACTCAA 286
QY 86 TGTCTTAGTTCCCTTCCCCAGAAACGACACGGAGCTGGCCA 131
Db 287 GAAACTCTTTCACAGGCTGGACAGAAACATCAGTCGCCCTGTCCA 332

RESULT 6
US-07-932-454A-2/c
; Sequence 2, Application US/07932454A
; Patent No. 5262318
; GENERAL INFORMATION:
; APPLICANT: GUTHRIE, ELLEN P.
; TITLE OF INVENTION: ISOLATED DNA ENCODING THE SphI
; TITLE OF INVENTION: RESTRICTION ENDONUCLEASE AND RELATED METHODS FOR
; TITLE OF INVENTION: THE SAME
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DAVID G. CONLIN; DIKE, BRONSTEIN, ROBERTS &
; ADDRESSEE: CUSHMAN
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: US
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/932,454A
; FILING DATE: 19920820
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAMS, GREGORY D.
; REGISTRATION NUMBER: 30901
; REFERENCE/DOCKET NUMBER: 42078

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 523-3400
; TELEFAX: (617) 5523-6440
; TELEX: 200291 STRE UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2692 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 703..1653
; OTHER INFORMATION: /note= "METHYLASE GENE STARTS AT
; POSITION 703/ENDS AT 1653. RESTRICTION
; OTHER INFORMATION: ENDONUCLEASE STARTS AT POSITION 1703/ENDS AT 2410"
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1703..2410
; US-07-932-454A-2

Query Match 12.4% Score 28.6; DB 1; Length 2692;
Best Local Similarity 55.6% Pred. No. 1.2; Gaps 0;
Matches 55; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

Qy 53 GTGGAAGCAGCGTCTGATTCGATGAAGTTTCAGTGTCTTCCTTCCTCCCGCAGAAAA 112
Db 418 GGGGAGGTTTCCACCGATTCCATGCATGATCTGATGCATGTCCCTTCCTCCAGTAGCA 359

Qy 113 ACGACAGCGGAGTGGCGCAAGACCTCTCTGGGTGATGAG 151
Db 358 GCGGTACGAGCGAGCAGACCCCGCCCGGACGGCTGGG 320

RESULT 7
US-08-766-354A-29
; Sequence 29, Application US/08766354A
; Patent No. 6013487
; GENERAL INFORMATION:
; APPLICANT: MITCHELL, LLOYD G.
; TITLE OF INVENTION: THERAPEUTIC MOLECULES GENERATED BY
; TITLE OF INVENTION: TRANS-SPLICING
; NUMBER OF SEQUENCES: 33
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKALDO, MARMELSTEIN, MURRAY, AND GRAM
; STREET: 655 15TH STREET, N.W., G STREET LOBBY, SUITE
; STREET: 330
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/766,354A
; FILING DATE: 13-DEC-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: KLESNER, SHARON
; REGISTRATION NUMBER: 36,335
; REFERENCE/DOCKET NUMBER: P8159-6006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1665 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```



```

; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: Concurrently Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/042,427
; FILING DATE: 27-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,205
; FILING DATE: 27-JAN-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/035,920
; FILING DATE: 22-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Hibler, David W.
; REGISTRATION NUMBER: 41,071
; REFERENCE/DOCKET NUMBER: UTSD:536
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 512/418-3000
; TELEFAX: 512/474-7577
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13865 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-09-009-217-11

Query Match 11.4%; Score 26.4; DB 3; Length 13865;
Best Local Similarity 55.4%; Pred.No.15;
Matches 51; Conservative 0; Mismatches 41; Indels 0; Gaps

QY 23 GCACAGAGATAGTCTGCTGTGGCATCCCTCGAAGGCACGTCGATTCGATGAGGTT 82
Db 7317 GCACATTGATGATGAATAACTTCTCTGTGAAGGCCAGCTGTTAGAATGAAGTG 7376
QY 83 CAGTGTCTAGTTCCTCCTCCCGCAGGAAAAAC 114
Db 7377 TTGTGTGAGTGTTAGGCCCGCAGGAGAGAC 7408

RESULT 14
US-09-009-656-11
; Sequence 11, Application US/09009656
; Patent No. 6132730
; GENERAL INFORMATION:
; APPLICANT: Thorpe, Phillip E.
; APPLICANT: King, Steven W.
; APPLICANT: Gao, Boming
; TITLE OF INVENTION: COMBINED TISSUE FACTOR AND FACTOR VIII
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR COAGULATION AND TUMOR
; TITLE OF INVENTION: TREATMENT
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/009,656

```


; PRIOR APPLICATION DATA:

THIS PAGE BLANK (USPTO)